



Title	応用数理サマーセミナー「逆問題」
Author(s)	Sakajo, Takashi
Citation	Hokkaido University technical report series in mathematics, 100, 1-190
Issue Date	2005-01-01
DOI	10.14943/626
Doc URL	<a href="http://hdl.handle.net/2115/684">http://hdl.handle.net/2115/684</a> ; <a href="http://eprints3.math.sci.hokudai.ac.jp/1059/">http://eprints3.math.sci.hokudai.ac.jp/1059/</a>
Type	bulletin (article)
File Information	tech100.pdf



[Instructions for use](#)

# 応用数理サマーセミナー「逆問題」

日本応用数理学会・北大COE共催

事業委員：速水 謙、奈良 高明、降旗 大介、  
松尾 宇泰、櫻井 鉄也、坂上 貴之

Series #100. August, 2005

**HOKKAIDO UNIVERSITY**  
**TECHNICAL REPORT SERIES IN MATHEMATICS**

- #74 T. Ozawa, Y. Giga, S. Jimbo and G. Nakamura (Eds.), Partial Differential Equations, 51 pages. 2002.
- #75 D. Matsushita (Ed.), Proceedings of the workshop “Hodge Theory and Algebraic Geometry”, 191 pages. 2003.
- #76 M. Hayashi and G. Ishikawa (Eds.), 2002 年度談話会・特別講演アブストラクト集, 34 pages. 2003.
- #77 T. Ozawa, Y. Giga, S. Jimbo, K. Tsutaya, Y. Tonegawa and G. Nakamura(Eds.), Proceedings of the 28th Sapporo Symposium on Partial Differential Equations, 76 pages. 2003.
- #78 S. Izumiya, G. Ishikawa, T. Sano and I. Shimada (Eds.), The 12th MSJ-IRI “Singularity Theory and Its Applications” ABSTRACTS, 291 pages. 2003.
- #79 H. Kubo and T. Ozawa (Eds.), Proceedings of Sapporo Guest House Symposium on Mathematics 15 “Evolution Equations”, 31 pages. 2003.
- #80 S. Miyajima, F. Takeo and T. Nakazi (Eds.), 第 12 回関数空間セミナー報告集, 122 pages. 2004.
- #81 Y. Giga, S. Izumiya and K. Deguchi (Eds.), Mathematical Aspects of Image Processing and Computer Vision 2003, 48 pages. 2004.
- #82 I. Shimada and Y. Tonegawa (Eds.), 2003 年度談話会・特別講演アブストラクト集, 52 pages. 2004.
- #83 The 2nd HU and SNU Symposium on Mathematics Abstracts, 22 pages. 2004.
- #84 T. Ozawa, Y. Giga, S. Jimbo, G. Nakamura, Y. Tonegawa and K. Tsutaya (Eds.), Proceedings of the 29th Sapporo Symposium on Partial Differential Equations, 77 pages. 2004.
- #85 T. Ozawa and Y. Tsutsumi (Eds.), Lectures on nonlinear dispersive equations I, 147 pages. 2004.
- #86 T. Ozawa and Y. Tsutsumi (Eds.), Lectures on nonlinear dispersive equations II, 47 pages. 2004.
- #87 T. Ozawa and Y. Tsutsumi (Eds.), COE Symposium Nonlinear Dispersive Equations, 85 pages. 2004.
- #88 T. Namiki, M. Hatakeyama, S. Tadokoro and H. Aoi (Eds.), 北海道大学数学教室におけるメタデータ交換プロトコル OAI-PMH に準拠した e-print サーバ構築, 14 pages. 2004.
- #89 S. Izumiya (Ed.) M. Takahashi, T. Miyao, G. Okuyama, Y. Nakano and K. Inui, 第 1 回数学総合若手研究集会 COE Conference for Young Researchers, 143 pages. 2005.
- #90 J. Saal, 1st COE Lecture Series  $H^\infty$ -calculus for the Stokes operator on  $L_q$ -spaces, 34 pages. 2005.
- #91 S. Miyajima, F. Takeo and T. Nakazi (Eds.), 第 13 回関数空間セミナー報告集, 111 pages. 2005.
- #92 N. Umeda, 第 4 回 COE 研究員連続講演会 反応-拡散方程式の大域解と爆発解について, 8 pages. 2005.
- #93 K. Arima, 第 2 回 COE 研究員連続講演会 極小モデルプログラムの入門およびその正標数への拡張, 25 pages. 2005.
- #94 Y. Nakano, 学位論文 Doctoral thesis “OPTIMAL HEDGING IN THE PRESENCE OF SHORTFALL RISK” 43 pages. 2005.
- #95 Keiji Matsumoto and Masao Jinzenji (Eds.), 2004 年度談話会・特別講演アブストラクト集, 17 pages. 2005.
- #96 T. Ozawa, Y. Giga, S. Jimbo, G. Nakamura, Y. Tonegawa and K. Tsutaya (Eds.), Proceedings of the 30th Sapporo Symposium on Partial Differential Equations, 83 pages. 2005.
- #97 M. Watanabe, 第 5 回 COE 研究員連続講演会 『逆散乱法』入門, 52 pages. 2005.
- #98 M. Takeda, T. Mikami (Eds.), Probability and PDE, 48 pages. 2005.
- #99 M. Van Manen, The 6th COE Lecture Series “From the cut-locus via medial axis to the Voronoi diagram and back” 42 pages. 2005.

応用数理サマーセミナー  
逆問題

札幌, 2005

日本応用数理学会・北海道大学数学 COE 共催



## 序

本冊子は、応用数理サマーセミナー「逆問題」(日本応用数理学会・北海道大学数学 COE 共催, 2005 年 9 月 5 日～7 日, 北海道大学) のレクチャーノートとしてまとめられたものである。

応用数理の中核を成す一分野であると同時に, 工学に遍在する「逆問題」に関して, その数理的基礎, 統計的手法から, 工学における応用, 最先端の話題, 数値計算法とソフトウェア実装にいたるまで, この分野を第一線で牽引されている講師陣に, 初歩的な所から詳しく解説していただいている。本セミナーが, 逆問題の基礎の習得と同時に, 応用数理に関連する若手研究者の交流の場となれば幸いである。

組織委員: 奈良 高明, 松尾 宇泰, 降旗 大介, 櫻井 鉄也, 坂上 貴之, 速水 謙

## プログラム

9/5(月)

13:00-15:45 山本 昌宏 (東京大学) 「基礎Ⅰ：応用逆問題と数学解析」

16:00-18:45 田邊 國士 (早稲田大学) 「基礎Ⅱ：統計的視点からの逆問題」

9/6(火)

9:00-11:45 久保 司郎 (大阪大学) 「応用：工学分野等における逆問題の展開」

13:00-15:45 池畠 優 (群馬大学) 「先端的話題：逆問題の解の直接再構成」

16:00-18:30 サマーセミナー参加者によるポスター発表

9/7(水)

9:00-11:45 藤原 宏志 (京都大学) 「数値計算：多倍長計算の逆問題、非適切問題への適用」



# 不適切問題への 統計的アプローチ

情報量規準

田 辺 國 士

## § 1. 不適切問題とモデル化

$x$  軸に沿って無限に長い棒があり、その側面が断熱されているとする。時刻  $t$  におけるこの棒上の位置  $x$  の温度を  $u(x, t)$  とすると、棒の温度分布  $u(x, t)$  の時間変化は、熱(拡散)方程式

$$\partial u / \partial t = \partial^2 u / \partial x^2 \quad (t > 0) \quad (1)$$

に従う。たとえば、 $t=0.16$  において図1の点線のように温度分布しているとき、 $t=1$  での分布は実線になる。これらは  $t=0$  で  $x=\pm 1$  のところに等しい熱を与えたときに実現する。

ここで  $t=1$  での分布を知って  $t=0.16$  での分布がどうなっていたかを求める問題を考えよう。ある温度分布を与え、それ以後の時点での分布を問う問題が直接問題とよばれるのに対して、上のような問題は逆問題とよばれる。 $0 < t_1 < t_2$  のとき、これらの時点での温度分布の間には

$$\int_{-\infty}^{\infty} (1/\sqrt{4\pi(t_2-t_1)}) \exp(-(x-\xi)^2/4(t_2-t_1)) u(\xi, t_1) d\xi = u(x, t_2) \quad (2)$$

という関係があるので、この逆問題は、データ  $u(x, t_2)$  を与えて第1種の Fredholm 型積分方程式(2)を解くことに帰着する。ところが式(2)や図1からわかるように、熱の拡散には強い平滑化作用があり、温度分布の空間的な振動は時間が経過すると共に急速に減衰する。したがって、ある時点で非常に異なる2つの温度分布でも、時間が経過するにつれ非常に似かよった分布になることがある。いいかえると、方程式(2)の右辺の小さな誤差が、逆問題の解に増減的に作用して、元の解とは全く異なる様相を示すものとなったり、さらには解が全



「ヒッチア」マンチオフタ

く存在しなくなりうる。このような意味で、この逆問題は不適切問題 *incorrectly posed problem* とよばれる。

不適切問題に導かれるものには、このほか、電場の様子から電荷の分布を求める問題や、観測システムの出力信号から入力信号を求める問題のように、素朴な意味での因果関係にある二つの事象の結果の方から原因を探ぐるといふ型の逆問題が多い。これらは、実用上多くの重要な問題を含み、しばしば第一種の積分方程式

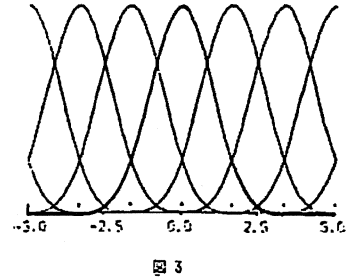
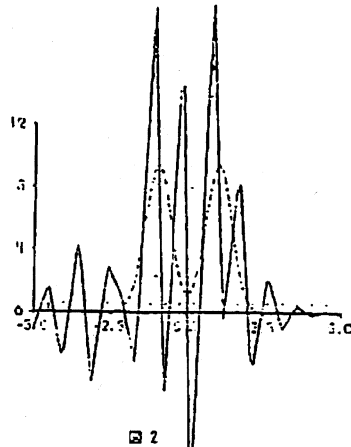
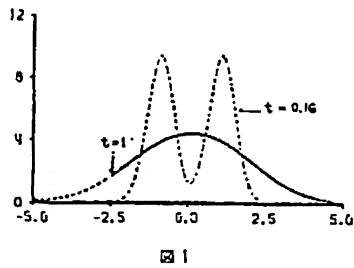
$$(Kf)(x) \equiv \int k(x, \xi) f(\xi) d\xi = g(x) \quad (3)$$

を解くことに帰着する。たとえば、走査型分光光度計などから得られる出力スペクトル  $g(x)$  と装置分解能関数  $k(x, \xi) \equiv k(x-\xi)$  および入力スペクトル  $f(\xi)$  とは(3)の関係にあると考えられるから、前二者を知れば、方程式(3)を解くこと(deconvolution)により分解能の改善が可能になる。また、核融合の実験においてプラズマの断面密度を測定する際には、第1種の Abel 型積分方程式

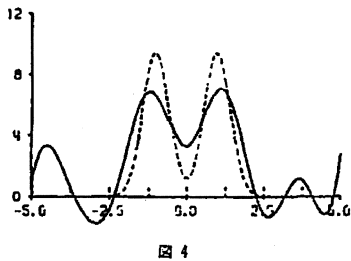
$$\int_0^x (1/\sqrt{x-\xi}) f(\xi) d\xi = g(x), \quad (0 \leq x \leq a) \quad (4)$$

を解く問題が生じる。

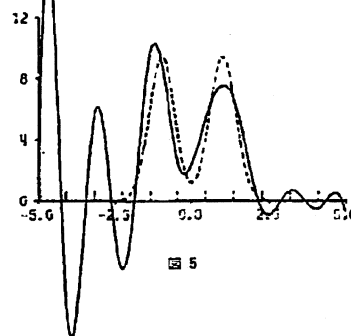
これらの第1種の積分方程式の解  $f$  の右辺  $g$  にたいする対応は、さきの例のように不連続なので、その数値解を得るために方程式(3)を直接離散化すると、導かれる連立一次方程式は悪条件となる。したがってそれを直接解いても右辺  $g$  にもともと含まれていた誤差や離散化誤差のために、安定した意味ある解が得られない。熱方程式の逆問題のときには、たとえば図2のようになる。これらの問題は文字どおり定式化が不適切な問



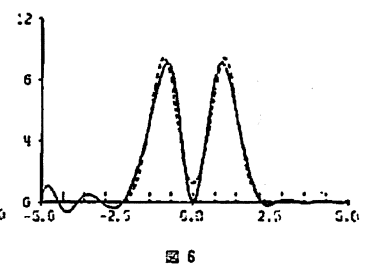
NSPL 3  
POINT 7  
RESIDUAL 0.2011275 E-6  
AIC -3077.2644  
MSGDEV 1.6059665  
MAXDEV 3.3510408



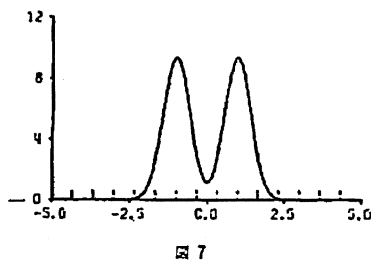
NSPL 3  
POINT 10  
RESIDUAL 0.1031005 E-10  
AIC -5047.1914  
MSGDEV 4.0487365  
MAXDEV 16.0337524



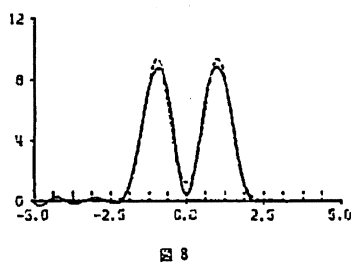
NSPL 3  
POINT 13  
RESIDUAL 0.3737510 E-17  
AIC -8020.2503  
MSGDEV 0.5063076  
MAXDEV 1.2355665



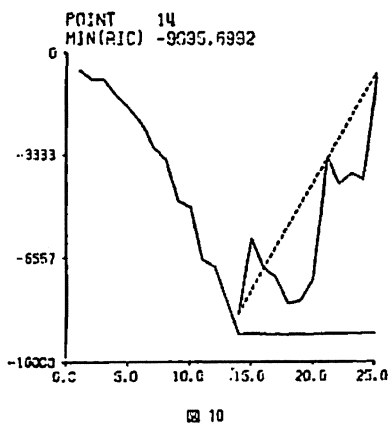
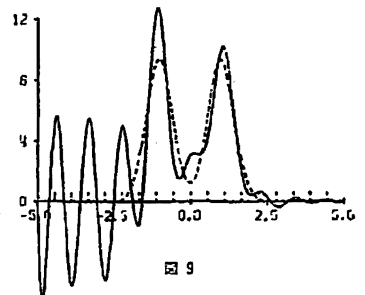
NSPL 3  
POINT 14  
RESIDUAL 0.1859564 E-13  
AIC -9035.6992  
MSGDEV 0.1709569 E-1  
MAXDEV 0.4757404 E-1



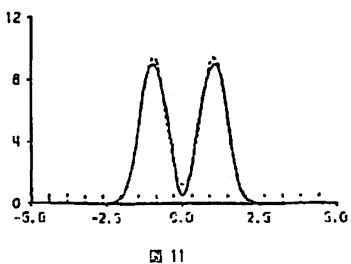
NSPL 3  
POINT 15  
RESIDUAL 0.1903763 E-10  
AIC -9688.9605  
MSGDEV 0.3065368  
MAXDEV 0.6472640



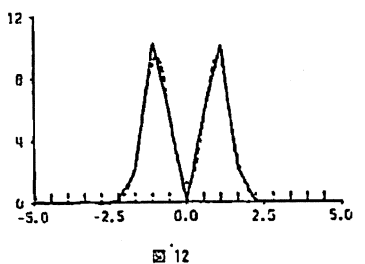
NSPL 3  
POINT 17  
RESIDUAL 0.1634534 E-13  
AIC -9032.4258  
MSGDEV 2.4102526  
MAXDEV 6.3556476



NSPL 2  
POINT 16  
RESIDUAL 0.1850769 E-13  
AIC -9033.5703  
MSGDEV 0.2323272  
MAXDEV 0.7056131



NSPL 1  
POINT 17  
RESIDUAL 0.1854156 E-19  
AIC -9094.2652  
MSGDEV 0.3029265  
MAXDEV 1.2773104



題なので、方程式(3)だけから近似解を見つけることは難しい。意味ある近似解を求めるためには、真の解あるいは誤差に関するなんらかの事前情報を用いて解の属すべき範囲を限定する必要がある。いいかえると、解にたいする広い意味でのモデルを導入し、問題を適切化しなければならない。この意味で、不適切問題はモデル化の問題といえることができる。

## § 2. 従来の適切化法

方程式(3)において解  $f$  がなめらかであることが予めわかっている場合がある。この知識を用いて、この方程式の妥当な近似解を得ることができる。

### A) Tikhonov の正則化法

$f$  のなめらかさを測る汎関数 (正則化汎関数)

$$\Omega(f) \equiv \sum_{i=1}^P \int_a^b \omega_i(x) (\partial^2 f(x) / \partial x^2)^2 dx \quad (5)$$

を考える。ただし、 $\omega_i(x)$  は正值連続関数とする。データ  $\bar{q}$  にたいして適当な正数  $\alpha$  を選び、汎関数 (平滑化汎関数)

$$M^\alpha(f, \bar{q}) \equiv \|Kf - \bar{q}\|_2^2 + \alpha \Omega(f), \quad (\|\cdot\|_2 \text{ は } L_2 \text{ ノルム}) \quad (6)$$

を最小にする  $\hat{f}$  を  $R_\alpha \bar{q}$  とする。このとき、 $Kf = 0 \Rightarrow \|f\|_2 = 0$  かつ  $\Omega(f) < \infty$  ならば、 $\bar{q} \in L^2$  にたいして  $R_\alpha \bar{q} \in C^p$  が一意的に定まり、 $\|\bar{q} - \bar{q}\|_2 = \delta$  が 0 に近づくとき、 $\alpha$  を  $\delta^2$  に比例するように小さく選んでいくと  $(R_\alpha \bar{q})(\xi)$  は  $f(\xi)$  に一様に近づく。この意味で、 $R_\alpha \bar{q}$  は  $f$  の近似解と看做することができる。(6) 式の中の  $\Omega$  は  $(R_\alpha \bar{q})(x)$  に振動的成分が現れるのを抑制する働きをする。 $R_\alpha, \alpha$  はそれぞれ正則化作用素、正則化パラメータとよばれる。

### B) 周波数空間でのモデル化による解法

$f$  を Fourier 変換した周波数空間で、 $f$  のなめらかさをモデル化することもできる。ここでは、方程式(3)は Convolution 型

$$(k \otimes f)(x) \equiv \int_{-\infty}^{\infty} k(x-\xi) f(\xi) d\xi = q(x) \quad (7)$$

とする。データ  $\bar{q}(x)$  が  $q(x)$  と '雑音'  $e(x)$  の和とすると、方程式(7)の形式解は

$$\hat{f}(\xi) \sim (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{i\omega\xi} (\hat{G}(\omega)/K(\omega)) d\omega \quad (8)$$

$$\|\hat{f} - f\|_2^2 \sim \int_{-\infty}^{\infty} |E(\omega)|^2 / |K(\omega)|^2 d\omega \quad (9)$$

となる。ただし、 $\hat{G}, K, E$  はそれぞれ  $\bar{q}, k, e$  の Fourier 変換とする。装置関数に対応する  $K(\omega)$  は一般に  $\omega \rightarrow \infty$  のとき  $|K(\omega)| \rightarrow 0$  となるのにたいし、雑音  $E$  は

一般に高い周波数成分をもつので、(8)(9)の右辺は意味を持たない。そこで雑音成分が優越する高周波部分を避けて、 $\hat{G}(\omega)/K(\omega)$  を有限の周波数帯  $|\omega| < \omega_0$  で反転させることにすれば、なめらかな近似解

$$\hat{f}(\xi) = (1/\sqrt{2\pi}) \int_{-\omega_0}^{\omega_0} e^{i\omega\xi} (\hat{G}(\omega)/K(\omega)) d\omega \quad (10)$$

が得られると考えられる。これを更に一般化して、 $|\omega| \rightarrow \infty$  のとき  $|S(\omega)| \rightarrow 0$  となる  $S(\omega)$  を選び、

$$\hat{f}(\xi) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{i\omega\xi} (\hat{G}(\omega)/K(\omega)) S(\omega) d\omega \quad (11)$$

を近似解とすることもできる。たとえば、

$$S(\omega) = K(-\omega)K(\omega) / (K(-\omega)K(\omega) + \alpha W(\omega)) \quad (12)$$

とすると、Tikhonov の正則化に対応する適切化が行なわれる。ただし、 $W(\omega)$  は非負の偶関数で、 $W(\omega) > 0$  ( $\omega \neq 0$ ),  $W(\omega) \geq k > 0$  ( $\omega \rightarrow \infty$ ) とする。 $S$  の Fourier 逆変換を  $s$  とすると、近似解(11)はデータ  $\bar{q}$  を  $s$  で平滑化した  $s \otimes \bar{q}$  を右辺とする方程式

$$Kf = s \otimes \bar{q} \quad (13)$$

の解に対応していることになる。

上記 A, B いずれの方法においても、近似解の平滑化の程度を決定する正則化パラメータ  $\alpha$ , 切り捨て周波数  $\omega_0$ , あるいはフィルター関数  $S$  の設定は、人間の主観的な判断に委ねられており、近似解の良し悪しは、全くこの判断の結果に依存する。

## § 3. やわらかい統計的モデル族の評価と AIC

不適切問題を解く鍵は、解  $f$  に関する既知の知識をいかに適切にモデルに表現するかにかかっている。データ外の知識には、漠然としたものから明確で限定的なものまで様々なレベルがあり、指定されるモデルもそれに応じて、現象論的なものから実体的なものまで多種多様となる。問題を不適切ならしめているものが誤差であり、それが観測や離散化に起因し制御しきれない性格のものであるとすれば、ここに統計的モデルを想定することは、きわめて自然であろう。

解  $f$  の構造がかなり知られており、それを一意的にモデルに表現できる場合には、問題は単一のモデルに含まれるパラメータの推定だけということになる。さきの熱方程式の逆問題の場合に、 $t=0$  で棒上の 2 点に熱が加えられたことを知れば、 $u(x, t)$  のモデルとして

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} \{ A \exp(-(x-B)^2/4t) + C \exp(-(x-D)^2/4t) \} \quad (14)$$

が考えられる。これを(2)式の左辺に代入して最小 2 乗法によってパラメータ  $A, B, C, D$  を決めれば、

$u(x, 1)$  の良い近似解が得られよう。

しかし、一般に実験科学では、対象に関する不完全な知識に基づいて有限のデータから解  $f$  の構造を帰納するという過程を繰り返しながら、漸次「真」の構造に近づいていく。その中途段階で用いられるモデルは単一のものではなく、対象の構造と誤差を分離できる程度には限定的であるが、同時に、データに柔軟に追従できる程度には十分可塑的でやわらかいモデルの族でなければならない。ここに、このモデルの族の中からデータを最も良く説明するものを評価選択するという問題が生じる。実験帰納においてモデル作成とならぶ重要な柱であるモデルの比較評価という仕事は、これまでデータ解析者の知識・経験・直感に頼って行なわれてきた。これをデータによって行なおうとすると、本質的な困難におつかる。やわらかいモデルの族において、自由度（パラメータ数）が比較的大きいモデルは、データへの見かけ上のあてはまりは良くなるが、データに含まれる誤差に過剰に反応してしまうため、安定した構造情報を取り出すことができない。逆に自由度が小さすぎるモデルは、対象の構造を十分反映できなくなり、得られる情報に偏りを生じる〔7〕。したがって、限られた数のデータによってモデルを評価選択するためには、この「安定性」と「偏り」を妥協させる「客観的」なモデルの評価基準が必要となる。

赤池氏は統計的モデルの選択の問題を、モデルが定める確率分布の「真」の確率分布への近似度の評価の問題として捉え、モデルの評価基準

$$\begin{aligned} \text{AIC} = & -2 \log_e (\text{モデルの最大尤度}) \\ & + 2 (\text{モデルの自由度}) \end{aligned} \quad (15)$$

を提案している〔1, 2, 3〕。モデルと真の分布の間の Kullback-Leibler 情報量を近似の良さの尺度の規範として採用し、最尤推定量に基づくこの情報量の推定量としてこの統計量を導びき、これを最小とするモデルを最良のものと判断する（AIC 最小化推定法）。この方法は、人間の判断に全面的に依存する従来の方法と異なり、データに基づくモデルの評価が可能になる点でユニークなものである。この AIC を適切に利用することによって、不適切問題の実用的な解法が得られることを次に示そう。

#### § 4. スプライン関数モデルによる不適切問題の解法

ここでは、データ  $\bar{q}(x)$  が離散点  $\{x_i\}$ ,  $(i=1, 2, \dots, M)$  上で与えられているとする。

解  $f$  に関する知見にもとづき、基底関数系  $\{\varphi_j(\xi)\}$ ,  $(j=1, 2, \dots, N)$  を選び、 $f$  を  $\varphi_j$  の一次結合で近似する。

すなわち、 $f$  を

$$\hat{f}(\xi) = \sum_{j=1}^N (\alpha_j / \lambda_j) \varphi_j(\xi) \quad (16)$$

とパラメトライズする。ただし、 $\lambda_j$  は、 $K\varphi_j$  を適当に正規化したものを  $\phi_j \equiv \lambda_j^{-1} K\varphi_j$  とするときの正規化係数とする。ここで、統計的線型モデル

$$\begin{aligned} \bar{q}(x_i) &= (K\hat{f})(x_i) + \varepsilon_i \\ &= \sum_{j=1}^N \alpha_j \phi_j(x_i) + \varepsilon_i, \quad i=1, 2, \dots, M \end{aligned} \quad (17)$$

を想定する。ただし、 $\varepsilon_i$  は平均 0、未知分散  $\sigma^2$  の正規分布に独立に従うものとする。積分方程式 (3) は、 $M \times N$  行列  $(\phi_j(x_i))$  を係数行列とるす連立一次方程式に離散化され、問題は  $(\alpha_j)$  を求めることに帰着する。 $(\alpha_j)$  の最小二乗解を計算すれば、 $f$  の近似解  $\hat{f}$  が (16) 式より求められる。このように、 $x$  空間上で回帰させたものを、 $\xi$  空間にもどす方法を仮りに随伴回帰法ともよび、関数系  $\{\varphi_j(\xi), \phi_j(x)\}$  を随伴回帰関数とよぶことにする。関数系  $\{\varphi_j\}$  の選択が、モデルの良否を決定する。

第 1 節の熱方程式の逆問題を解くのに、 $\{\varphi_j(\xi)\}$  として、一定の小さな区間のみに台をもつ関数  $\varphi_j$  を要素とする  $B$ -スプライン関数系とよばれるやわらかいモデル族を用いる。たとえば図 3 のように、3 次のスプライン基底関数  $\varphi_j$  は、相隣り合う 4 つの区間の外では 0 となり、各区間上では 3 次の多項式で、区分点でなめらかに接続している。 $f$  の考察の対象区間を  $[-5, 5]$  に限定し、この区間内に POINT 個の区分点を配するとき、NSPL 次のスプライン関数族  $\{\varphi_j\}$  の基底は  $N = \text{NSPL} + \text{POINT} + 1$  個の要素で構成され、次数が高いほどこの区間でなめらかな関数族を式 (16) により表現する。区間  $[-2.5, 5]$  上に一様に撒かれた  $M = 201$  個の点上で関数  $u(x, 1)$  を計算し（図 1 の実線部）、平均 0、標準偏差  $10^{-11}$  の正規乱数を加えたものをデータ  $\bar{q}$  として、上記の方法で解いた近似解が図 4—9, 11, 12 に示されている。図 4—9 は 3 次のスプライン関数を用い、その数すなわち区分点（くしのはのような短い線で示している）を順次ふやして  $f$  を推定したものである。区分点の数が少ないと解をよく近似できないし、逆に多すぎると誤差の影響をうけることがわかる。図 10 はスプライン関数の数をふやしていくときの AIC の動きが示されている。右上方はその一部を拡大したもので、点線は連立一次方程式の残差平方和が一定のところを示す。図より内部区分点の数 (POINT) が 14 のとき (図 8), AIC が最小であり、推定された  $\hat{f}$  が真の解 (点線) とよく一致していることが観察できる。図 11, 12 には同じデータに

たいするそれぞれ二次と一次のスプライン関数によるモデル族の中で、それぞれ AIC が最小となるものを示している。この二つの AIC の値は図 7 のそれよりも大きい。AIC を比較することにより、異なる次数のスプライン関数によるモデル族の良さを比べることができる。なお図中、RESIDUAL は連立一次方程式の残差平方和、AIC は統計量  $AIC = M \log(\text{RESIDUAL}) + 2(\text{NSPL} + \text{POINT} + 1)$ 、MSQDEV は真の解と推定関数との差の、区間  $[-5, 5]$  内 301 点上の平均 2 乗偏差の平方根、MAXDEV は同じく最大偏差を示している。AIC を用いることによりこのようなやわらかいモデルの族の中からデータにもとづいて、良い近似解を見出すことができることは、非常に興味深い。ここでは区分点を等間隔にとるモデルを述べたが、スプラインの区分点の数だけでなく、その位置をも変数とする非線型モデルを、人工的誤差を加えない実際の問題に適用して良好な結果を得ている。また Abel 型方程式 (4) も、随伴回帰関数系として  $\{\xi^{1/2}, \eta\}$  を選ぶことにより、上と同様に随伴回帰法に AIC を用いて、かなり広い問題に対して良好な結果を得ている。また、離散化された問題を特異値分解法によって適切化することもできるが、これも一種の随伴回帰法と考えることができる [5, 6]。

## § 5. む す び

解析者が用意するモデルの族が適切で、真の構造を十

分反映するモデルを含んでいる場合には、AIC 最小化法はかなり良好に作動する。良いモデル族を作るのはあくまでも人間であるが、モデル評価という人間が行なう帰納推理の本質的な仕事の一部を、機械的方法によって代行することを可能にしたという意味で、AIC 最小化法は、いわば人工知能への一つのステップを示していると言えよう。パタン認識、画像処理、リモートセンシング等の領域への応用が期待される。

また、AIC 最小化法は、いわゆる簡潔性の原理の一表現と解釈することができるが、これが予測という立場から導かれた点は興味深い。

## 参 考 文 献

- [1] H. Akaike; Res. Memo No. 44, Inst. Statist. Math., 1971.
- [2] H. Akaike; Proc. 5th Hawaii Intl. Conf. Syst. Sci., pp. 99-101, 1972.
- [3] H. Akaike; Problems of Control and Information, AKADEMLAI KIADO pp. 267-281, 1973.
- [4] K. Tanabe; Res. Memo No. 62, Inst. Statist. Math., 1974.
- [5] K. Tanabe; *Computation and Analysis*, 6-4, pp. 2-25, 1975.
- [6] 田辺國士, 日科技連, 計算機活用セミナーテキスト pp. 41-62, 1975.
- [7] 田辺國士, *bit* 臨時増刊, 数値計算における誤差, pp. 113-125, 1975.

(たなべ・くにお, 統計数理研究所)



§ Proceedings  
§ of the

§ COMPUTER SCIENCE

§ AND  
§ STATISTICS:

§ ELEVENTH ANNUAL

§ SYMPOSIUM

§ ON THE  
§ INTERFACE

§ March 6 and 7,

§ 1978

§ Institute of Statistics

§ North Carolina State University

# Least Squares Spline for Incorrectly-Posed Problems: Information Theoretic Approach

Kunio Tanabe

Institute of Statistical Mathematics, Tokyo, Japan  
and  
Brookhaven National Laboratory, Upton, N.Y. 11973

The use of the least squares splines for solving ill-posed (multi-collinear) problems is discussed. It is shown that choosing appropriate number of nodes, locations of nodes and degree of splines, we can control the smoothness of approximation in terms of spline functions, without resorting to Tikhonov's regularization (ridge regression) method which is designed to adjust the smoothness of the model functions to be fitted to data. It is also demonstrated that the Akaike's Information Criterion can successfully be applied to select the appropriate number of free parameters in the 'spline regression' models in which the number of nodes, locations of nodes and degree of splines are variable, based solely on data without using a subjective criterion such as significance levels in the statistical hypothesis testing.

The recovery of some definite structures from noisy data forms an important subject in experimental sciences. The 'inverse problem' in physical sciences are typical sources of this type of problems. These problems are inherently ill-conditioned in the sense that small disturbances in data cause a disastrous effect on the direct estimate of the true structure, although sometimes the problems are considered improperly formulated. They are called 'incorrectly-posed problems' and extensive efforts have been devoted to them in the literature for a reasonable estimate of the structures.

The author re-examined the subject from the information theoretical point of view which was originally introduced by Akaike in his paper for fitting statistical models. Making use of the Kullback-Leibler information quantity and the asymptotic theory of likelihood ratio statistics, he proposed a new statistic based on what we call the Akaike's information criterion(AIC),

$$\begin{aligned} \text{AIC} = & -2\log_2(\text{maximum likelihood of a model with} \\ & \text{respect to given data}) \\ & + (\text{degree of freedom of the model}), \end{aligned}$$

and the model chosen is the one with the smallest AIC

value among competitive models with varying degree of freedom. This procedure which we call minimum AIC estimation(MAICE) differs from the existing methods in that it does not require such a subjective criterion as the significance level in statistical hypothesis testing. Although its significance in the field of time series analysis was well demonstrated in Akaike's papers [1], apparently its potential use in other areas such as pattern recognition is not yet fully recognized. In this note we show its applicability to the resolution of noisy data by treating two typical problems.

The first problem is the regularization of a noisy ill-conditioned system of linear equations,

$$Ax = b,$$

whose coefficient matrix and the right hand side constant vector are corrupted by random noise. The problem is closely associated with the numerical solution of the Fredholm integral equation of the first kind,

$$\int K(s,t)x(t)dt = b(s),$$

which arises in many branches of physical sciences,

typically in the experimental sciences where physical data are measured by indirect sensing devices.

As was well demonstrated by Phillips [6], direct application of usual numerical procedures to the equation will result in an oscillatory solution which is often not in good agreement with the nature of the problem. This subject has been discussed extensively in the literature and various smoothing procedures have been developed. They are classified into two categories. The first one minimizes

$$\|b - Ax\|^2 + \rho \|x - x_0\|_M^2,$$

where  $\|x\|^2 = x^*x$ ,  $\|x\|_M^2 = x^*Mx$  and the positive definite matrix  $M$ , the vector  $x_0$  and the smoothing parameter  $\rho$  are determined from a priori information on the 'true' solution. The result thus obtained is given by

$$x = (A^*A + \rho M)^{-1}(A^*b + \rho Mx_0).$$

Phillips [6] and Tikhonov [9] devised this method. Foster [3] and Strand & Westwater [7] gave statistical extensions and justifications of this approach to certain models in which disturbances in the coefficient matrix were not taken into account. Their method for determining the smoothing parameter is not entirely satisfactory since it depends on the rather precise knowledge of the covariance matrix of the statistical noise which is usually unknown. The other method uses the singular value decomposition to invert the equation by the least squares method for the rank deficient case. Golub & Saunders [4-a], Hanson [5] and Varah [11] used an ad hoc method for determining an effective rank of  $A$ . Both approaches will give satisfactory results so long as  $M$ ,  $x_0$  and  $\rho$  or an effective rank of  $A$  are suitably chosen. However the choices depend heavily on subjective judgement of the investigator.

To solve this problem we introduced a statistical model which uses spline functions (Fig. 1-3). Applying the MAICE procedure we can determine 'objectively' a reasonable approximate solution in which the gain in resolution is balanced against the amplification of the noise. Fig. 1-4~9 and 11~12 show the approximate (spline function) solutions of the 'backward heat equation for the system,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < +\infty.$$

This problem is reduced to solve the Fredholm integral equation of the first kind,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi(t_2-t_1)}} \exp\left(-\frac{(x-y)^2}{4(t_2-t_1)}\right) u(y, t_1) dy = u(x, t_2), \quad (0 < t_1 < t_2),$$

for a given  $u(x, t_2)$ . Approximating  $u(x, t_1)$  in terms of a finite linear combination of B-splines and discretizing  $x$ -space by collocation, we obtain a linear algebraic equation. Fig. 1-1 shows temperature profiles  $u(x, t)$  at the time  $t_1 = 0.16$  and  $t_2 = 1$ . Fig. 1-2 shows an solution obtained by applying the usual least squares method directly to the discretized linear equation. Fig. 1-3~9 show the least squares (cubic spline function) solutions with varying number of node points. Fig. 1-10 shows the profile of the values of AIC vs. number of nodes of cubic spline functions. Fig. 11 and 12 show quadratic and linear spline approximate solutions which are respectively selected by MAICE from quadratic and linear spline least squares solutions with various node points. Recently Golub [4-b] and Wahba [12] proposed methods to determine  $\rho$ , using similar statistical idea.

The second problem is the reconstruction of a polynomial regression from noisy measurement. The difficulty of this type of smoothing is due to our ignorance of degree of the polynomial to be fitted to given data. If we keep the order of the regression polynomial too low, the solution is biased from the true structure. If we allow too many parameters in the regression polynomial, the solution is affected with noise. Thus the main problem is the choice of a suitable degree of the regression polynomial. This problem was treated by Anderson [2] from the statistical hypothesis testing point of view. But his method is not satisfactory since it depend on subjective judgement. Applying the MAICE procedure, we are free from the limitation. Fig. 2 shows the 2-dimensional regression polynomials with degrees 0 to 10, fitted to data (top right) which are generated by adding random numbers to the polynomial (top left),

$$(2x^3 - x)(y^2 - y)/2$$

evaluated at grid points in  $x$ - $y$  plane.

The two problem sketched here will find direct application in the following important fields of science and engineering:

- a) super-resolution of noisy patterns respectively obtained by radio telescopes, by mine detectors, by electron scanning beams and X-ray photographs,
- b) deduction of temperature profiles in the

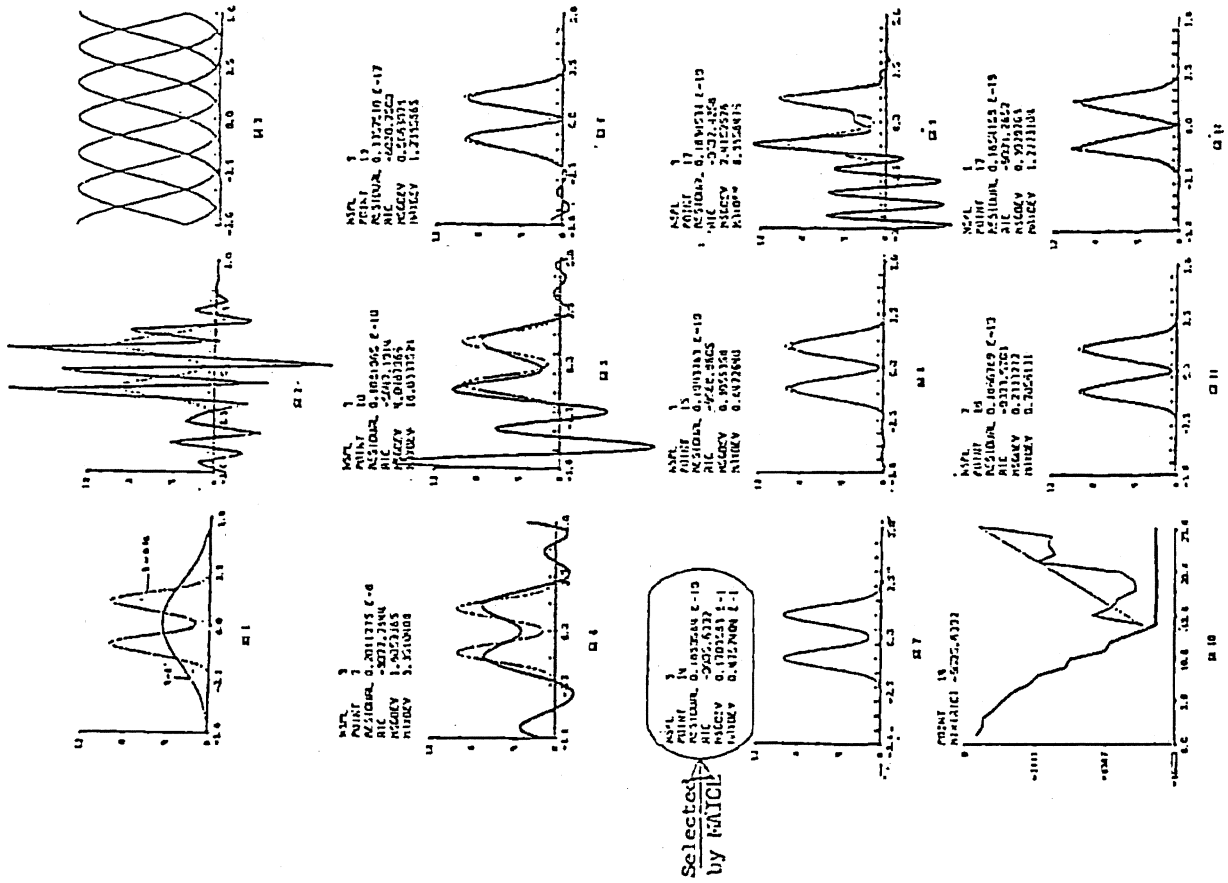


Fig.1 (K. Tanabe, Surikagaku, March 1976)

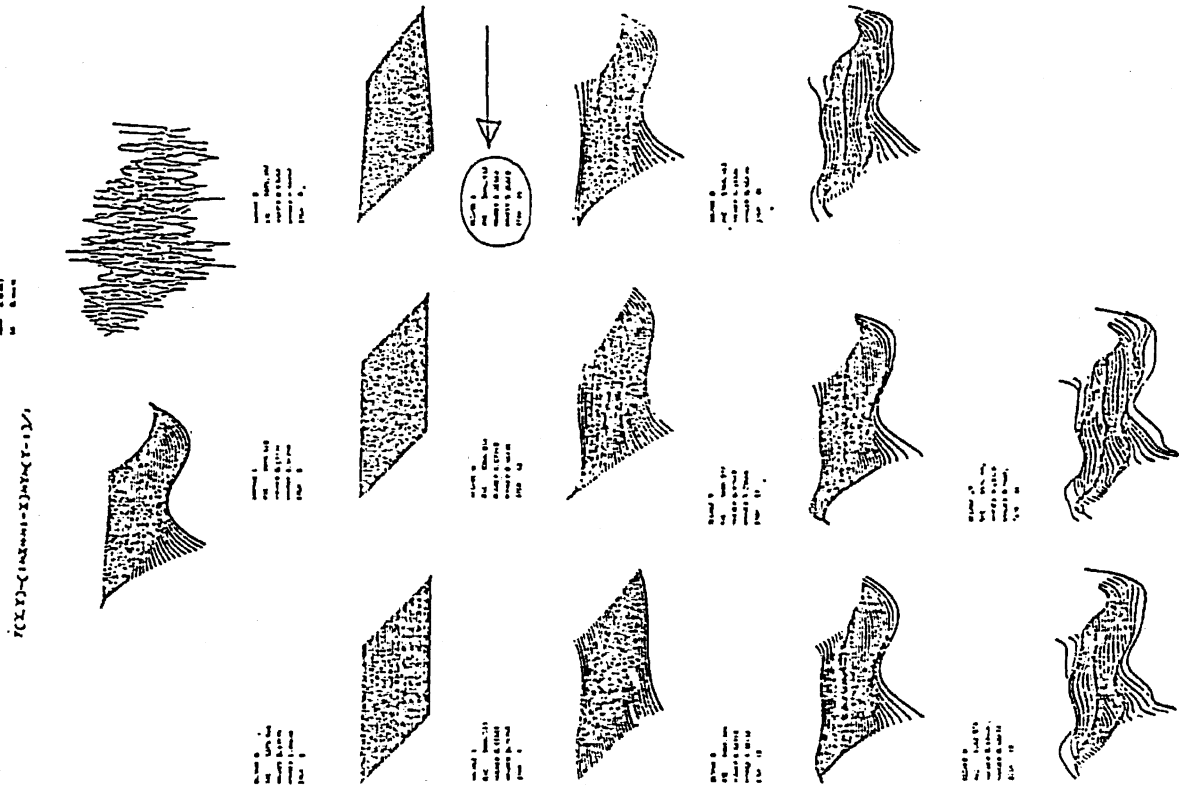


Fig.2 (K. Tanabe, 1974)

atmosphere by measuring the spectral distribution of infrared energy,

- c) inference of aerosol size distribution from diffusional decay measurements of the total particle counts,
- d) unsupervised estimation (learning without teacher) of mixture of density functions in pattern recognition,
- e) numerical solution of ill-posed problems such as harmonic continuation and numerical inversion of Laplace transform.

#### REFERENCE

1. a) H. Akaike, "Information theory and an extension of the maximum likelihood principle, 2nd International Symposium on Information Theory, Problems of Control and Information Theory, B.N. Petrov & F. Csaki eds., 1973, pp. 267-281.
  - b) H. Akaike, "Fitting autoregressive models for prediction" Ann. Inst. Statist. Math., vol. 21, pp. 243-247, 1969.
  - c) H. Akaike, "Statistical predictor identification" Ann. Inst. Statist. Math., vol. 22, pp. 203-217, 1970.
  - d) H. Akaike, "A new look at the statistical model identification" IEEE Trans. on Automatic Control, vol. AC-19, pp. 716-723, 1974.
  2. T.W. Anderson, "The choice of the degree of a polynomial regression as a multiple decision problem I" Ann. Math. Statistics, vol. 33, pp. 255-265, 1962.
  3. M. Foster, "An application of the Wiener - Kolmogorov smoothing theory to matrix inversion" SIAM Appl. Math., vol. 9, pp. 387-392, 1961.
  4. a) G.H. Golub and M.A. Saunders, "Linear least squares and quadratic programming" Integer and Nonlinear Programming, pp. 229-256, 1970.
  - b) G.H. Golub, M. Heath and G. Wahba, "Cross-validation and optimum ridge regression" Stanford Computer Science report, 1976.
  5. R.J. Hanson, "Integral equation of immunology" CACM, vol. 15, pp. 883-890, 1972.
  6. D.L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind" JACM, vol. 9, pp. 34-97, 1962.
  7. O.N. Strand and B.R. Westwater, "Minimum RMS estimation of the numerical solution of a Fredholm integral equation of the first kind" SIAM Numer. Anal., vol. 15, pp. 287-295, 1968.
  8. a) K. Tanabe, "Fitting regression curves and surfaces by Akaike's information criterion" Inst. Statist. Math. Res. Memo. no. 62, 1974.
  - b) K. Tanabe, "Statistical regularization of a noisy ill-conditioned system of linear equations by Akaike's information criterion, Computation and Analysis, vol. 6, pp. 2-25, 1975.
- Above two papers were presented at the 1974 IEEE International Symposium on Information Theory held at Univ. Notre Dame, Indiana, 1974.
- c) K. Tanabe, "Treatment of Statistical errors" (in Japanese), bit, vol. 7, pp. 113-125, 1975.
  - d) K. Tanabe, "Statistical Approach to Incorrectly Posed problems" (in Japanese), Surikagaku (Math. Sciences), no. 153, pp. 60-64, 1976.
  9. A.N. Tikhonov, "The stability of algorithms for the solution of degenerate system of linear algebraic equations" USSR Computational Math. and Mathematical Physics, vol. 5, pp. 181-188, 1965.
  10. Doney S., "The Application of numerical filtering to the solution of integral equations encountered in indirect sensing measurement" Jour. of Franklin Inst., vol. 279, pp. 95-109, 1965.
  11. J.M. Varah, "On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems, SIAM Numer. Anal., vol. 2, pp. 257-267, 1973.
  12. G. Wahba, "A survey of some smoothing problems and the method of generalized cross-validation for solving them" AMS Notes Short Course on Statistics, 1977.



## SYMPOSIUM

重力と地磁気の基礎(1) / 12

### ベイズモデルによる曲線・ 曲面のあてはめ

田 辺 國 士

たなべ くにお  
統計数理研究所

田 中 輝 雄

たなか てるお  
電気通信大学統計数理研究所

観測データを平滑化して構造情報を取り出す方法の1つに曲線・曲面のあてはめがある。この曲線・曲面の表現に離散スプラインがその可塑性の点を優れていることを示す。またそのなめらかさや相変化をベイズモデルに表現することにより、平滑度、異相の有無、相変化点の検出に赤池の情報量規準 ABIC が有効であることを示す。

#### 1. はじめに

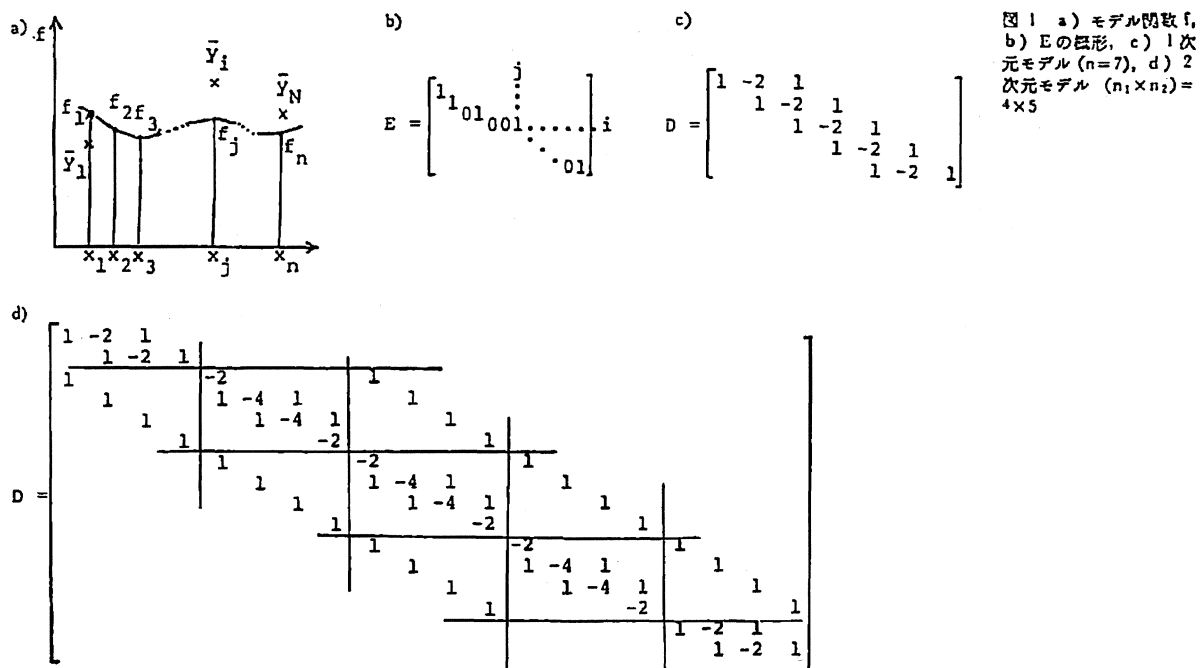
観測データから対象の構造に関する情報を引き出すことはあらゆる実験科学に共通の仕事である。データには偶然誤差がつきものであるから、図式的に

$$[\text{データ}] = [\text{構造}] + [\text{偶然誤差}]$$

と考えることができる。したがって、データの処理は右辺の2つの成分を分離することを意味する。しかしこの統計的処理が可能であるためには、「構造」と「偶然誤差」に関してなんらかの予想や事前情報がなければならない。データに先立つ情報に基づいて「構造」を数式モデルに表現し、構造モデルと偶然成分とのからみあい確率密度関数の形に表現すること（すなわち統計的モデリング）ができてはじめて両者の分離が可能となる。従来の統計学では、この2つの成分がパラメトリックモデルの形にアприオリに確定していることを前提に展開されているため、データをたくさん集めてパラメターの値を精密に推定することが統計的データ解析である、と理解されがちである。しかしデータ処理の主要な仕事はむしろ統計的モデリングにあり、パラメターの推定は副次的役割をもつにすぎないといえる。

統計的モデリングの基礎となる事前情報は漠然としたものであることが多く、指定されるべき統計的モデルも一意的には定まらない。したがって統計的モデリングにはかなり主観的要素が入り込む余地がある。統計的処理の客観性を希うあまり、主観的要素を排除してしまうならば統計学には何も残らない。この主観的要素こそが「構造」と「偶然誤差」の分離可能性の根拠であり、無から有は生じないのである<sup>[12-14]</sup>。しかしもちろん、主観性をむやみにふりまわすことは独善に陥ることが明らかである。データの統計的処理の成否は、モデリングにおける主観性をいかにうまく

原稿受理：1983年2月5日（編集部）



塩梅して、データの中に含まれる客観性を引き出すかにかかっている。統計的データ解析におけるこのような「方法的主観性」の役割の認識に立脚して、近年統計的モデリングにおけるベイズモデルの有用性が再評価されつつある<sup>[1-3,5-11,17]</sup>。本稿の目的は、曲面・曲線のあてはめを例にとり、データ処理のためのこの新しい方法を紹介することにある。

曲線・曲面のあてはめには多項式、三角関数やスプライン関数が「構造」を近似するモデルとしてしばしば用いられる。しかしこのような函数族を採用するということは、「構造」に関するかなり大きな主観的判断を下したことになる。なぜならこれらのパラメトリックなモデルはそれぞれ固有の偏りをもっており、データにあてはめたときに「構造」を反映しない近似函数の「くせ」が出てしまうからである。実際、多項式や三角関数で矩形波のように急激に変化する構造を表現しようするとギブス現象と呼ばれる不必要な振動が生じる。スプライン函数も節点の位置が適当でないと「構造」を忠実に再現することができない。また函数族に含まれる自由パラメータの数が多すぎるときにはデータに含まれる誤差に過剰に反応してしまうため、安定した構造情報を取り出すことができない。も

っともこの最後の点に関しては、赤池の情報量規準 AIC を用いて、妥当な自由度をデータに基づいて、決定することができる<sup>[12,14,15]</sup>。そこで本稿では近似函数の形状をパラメトリックモデルによって予め規定してしまわないで、対象の「構造」と「誤差」を分離しうる程には限定的であるが、同時に対象の構造に柔軟に追従できる可塑的でやわらかいモデルの族として「離散スプライン函数」を用いる方法を述べよう。

## 2. ベイズモデルと情報量規準 ABIC

推定すべき関数  $f(x)$  を、離散点  $x_j$  上の値  $f_j = f(x_j)$  で表現する。つまり、

$$f = (f_1, f_2, \dots, f_j, \dots, f_n)^t \quad (2)$$

と  $f$  をパラメトライズする ( $t$  は転置)。以下では関数とベクトル  $f$  を同一視する。  $x_j$  は等間隔にとり  $n$  は十分大きくとっておく。二次元問題では長方形領域を考え  $x_j$  を格子点 ( $n_1 \times n_2$ ) 上の点と考える ( $n_1 \times n_2 = n$ )。一次元問題の場合、図 1-a) のようになる。

今、  $N$  個のデータ  $\bar{y}_i (1 \leq i \leq N)$  が得られているとする。  $\bar{y}_1$  を、

$$\bar{y} = (\bar{y}_1, \bar{y}_2, \dots, \bar{y}_1, \dots, \bar{y}_N)^t \quad (3)$$

と表す。  $n$  は  $N$  より十分大きくとる。 fit-point  $x_j$  上

に常にデータ  $\bar{y}_1$  がある必要はない。

データ  $\bar{y}_1$  の誤差  $\varepsilon_1$  を

$$\varepsilon_1 \sim i.i.d.N(0, \sigma^2) \quad (4)$$

とすれば  $y_1$  の分布は、

$$p_1(y_1|f_j, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \times \exp\left(-\frac{1}{2\sigma^2}(y_1 - f_j)^2\right) \quad (5)$$

となる。ただし、 $\bar{y}_1$  は  $f_j$  上のデータである。このとき  $y$  の分布は、

$$p(y|f, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \times \exp\left(-\frac{1}{2\sigma^2}\|y - Ef\|^2\right) \quad (6)$$

と表せる。ここで  $\|\cdot\|$  は、ユークリッドノルムを表す。また  $E$  は、 $\bar{y}$  と  $f$  の対応を表す図 1-b) 行列である。

つぎに、 $f$  のなめらかさを二階差分の大きさを表現する。各点での二階差分の大きさを次のような量、

$$(\text{一次元モデル}) |f_{1,i-1} - 2f_{1,i} + f_{1,i+1}|, 2 \leq i \leq n_1 - 1 \quad (7)$$

$$(\text{二次元モデル}) |f_{1,j-1} + f_{1-1,j} - 4f_{1,j} + f_{1,j+1} + f_{1+1,j}|, \\ 2 \leq i \leq n_1 - 1, 2 \leq j \leq n_2 - 1, \quad (8)$$

$$|f_{1-1,j} - 2f_{1,j} + f_{1+1,j}|, \\ j = 1, n_2, 2 \leq i \leq n_1 - 1, \quad (9)$$

$$|f_{1,j-1} - 2f_{1,j} + f_{1,j+1}|, \\ i = 1, n_1, 2 \leq j \leq n_2 - 1, \quad (10)$$

で測り、 $f$  のなめらかさの程度を  $\|dDf\|^2$  で測る。 $D$  は構造のある(帯行列)スパースな行列である。各行の順はどのように取ってもかまわないが Givens 変換法による最小二乗計算の効率を上げるために図 1-c), d) のようにとる(3節参照)。

ここで、超パラメータ  $d \in \mathbb{R}^+$  をもつ  $f$  の improper な事前分布

$$\pi(f|d) = \left(\frac{1}{2\pi}\right)^{l/2} d^l \phi \exp\left(-\frac{1}{2}\|dDf\|^2\right) \quad (11)$$

を導入し、 $f$  のなめらかさをこの事前分布に含まれる超パラメータ  $d$  で制御する。一次元モデルの場合  $l = n - 2$ 、二次元モデルの場合  $l = n - 4$  で、 $\phi^2$  は  $D^t D$  の非零固有値の積である。ここでは improper な事前分布を用いているが、これはある種の proper な事前分

布を用いる方法の近似と考えられる<sup>[17]</sup>。

したがって、データ  $\bar{y}_1$  が与えられた時の周辺尤度  $L(\sigma^2, d)$  は

$$L(\sigma^2, d) = \int p(\bar{y}|f, \sigma^2) \pi(f|d) df \quad (12)$$

と定義される。

赤池はベイズ型情報規準量 ABIC<sup>[1,2,2]</sup>

$$ABIC = -2 \log L(\sigma^2, d) \quad (13)$$

を最小にする  $\sigma^2, d$  を  $f$  の推定に用いる方法を提案している。今、ABIC を最小にする  $\sigma^2, d$  を  $\hat{\sigma}^2, \hat{d}$  (最尤推定値)とすると、ベイズの定理により、 $f$  の事後分布は、

$$p(f|\bar{y}) = \frac{p(\bar{y}|f, \hat{\sigma}^2) \pi(f|\hat{d})}{\int p(\bar{y}|f, \hat{\sigma}^2) \pi(f|\hat{d}) df} \quad (14)$$

と表されるから、(14)式を最大にする  $f$  を最適近似関数  $\hat{f}$  として選ぶことにするのである。

$\alpha = d\sigma$  と超パラメータの変数変換を行なうと、

$$L(\sigma^2, \alpha) = \left(\frac{1}{2\pi}\right)^{N+l-n/2} \left(\frac{1}{\sigma}\right)^{N+l-n} \alpha^l \phi |\det(Z_\alpha^t Z_\alpha)|^{-1/2} \\ \exp\left(-\frac{1}{2\sigma^2}\|b - Z_\alpha f_*\|^2\right), \\ Z_\alpha = \begin{bmatrix} E \\ \alpha D \end{bmatrix}, b = \begin{bmatrix} \bar{y} \\ 0 \end{bmatrix} \quad (15)$$

ここで、 $\|b - Z_\alpha f_*\|^2$  は最小二乗残差で  $f_*$  はそのときの最小二乗解である。つまり、 $\hat{f}$  は  $\alpha = \hat{\alpha}, \sigma^2 = \hat{\sigma}^2$  のときの  $f_*$  に他ならない。よって、問題は(15)式を最大にする  $\sigma^2, \alpha$  を求めることになる。

$$\frac{\partial}{\partial \sigma^2} L(\sigma^2, \alpha) |_{\sigma^2 = \hat{\sigma}^2} = 0 \quad \text{より,}$$

$$\hat{\sigma}^2 = \frac{1}{N+l-n} \|b - Z_\alpha f_*\|^2 \quad (16)$$

これを(15)式に代入して、

$$L'(\alpha) = \left(\frac{1}{2\pi}\right)^{N+l-n/2} \phi \exp\left(-\frac{N+l-n}{2}\right) \alpha^l \\ |\det(Z_\alpha^t Z_\alpha)|^{-1/2} \\ \times \left(\frac{1}{N+l-n} \|b - Z_\alpha f_*\|^2\right)^{-(N+l-n/2)} \quad (17)$$

を得る。したがって、 $\alpha$  についてその尤度方程式

$$\frac{\partial}{\partial \alpha} L'(\alpha) |_{\alpha = \hat{\alpha}} = 0 \quad (18)$$

を解けばよいのであるが、(18)式は非線形方程式で解析的に解くことが困難である。そこで  $\alpha$  を数値的に求



める。いろいろな値 $\alpha$ に対して、

$$ABIC(\alpha) = (N+1-n) \log(\|b - Z_\alpha f_*\|^2) - 2 \log \alpha + \log |\det(Z_\alpha^* Z_\alpha)| + C \quad (19)$$

を計算し、この ABIC 最小になるような $\alpha$ を求め、この $\alpha$ に対応する $f_*$ を推定関数にする。(19)式の第3項を行列表項、第1項を残差項と呼ぶことにする。

### 3. Givens 変換法

ABIC( $\alpha$ )を計算するには、次の最小二乗問題

$$\min_f \|b - Z_\alpha f\|^2 \quad (20)$$

を解けばよい。いろいろな $\alpha$ に対して最小二乗関数を解く必要があるので解法は高速であることが必要である。最小二乗問題を解く場合、ふつう Householder 変換法が用いられるが、スパースな大行列に対しては Givens 変換法の方がより有効な解法である[10]。ここでも Givens 変換法を用いることにより、行列 $Z_\alpha$ の特殊な構造に着目して fill-in (零要素が非零要素になること)をおさえ、Householder 変換法より計算量を減らすことができることを示そう。

(20)式を解くにはまず $Z_\alpha$ を QR と分解する。Qは直交行列で、Rは $(r_{ij})$ とすれば $i > j$ の部分がすべて零である行列である。 $\|Q\| = \|Q^*\| = 1$ に注意すると、

$$\|b - Z_\alpha f\|^2 = \|Q^* b - R f\|^2 = \left\| \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} - \begin{bmatrix} R_1 \\ 0 \end{bmatrix} f \right\|^2 \quad (21)$$

であるから、

$$b_1 = R_1 f \quad (22)$$

となり、後退代入より $f$ が求まる。そのとき残差は $\|b_2\|^2$ となり(19)式に残差項が求まる。また(19)式を行列表項は、

$$\log |\det Z_\alpha^* Z_\alpha| = \log |\det R^* R| = \log \|r_{11}\|^2 = 2 \sum \log |r_{11}| \quad (23)$$

となり求めることができる。よって、ABIC( $\alpha$ )の値は $Z_\alpha$ を QR と分解することによって求めることができるわけである。このとき、変換行列Qとして二次元の回転行列を用いるのが、Givens 変換法である。

$z_{ik}$ を零にする変換行列 $Q_{ik}$ を図2-a)のように構成する。ここで、

$$r = \sqrt{z_{kk}^2 + z_{ik}^2}, c_{ik} = z_{kk}/r, s_{ik} = z_{ik}/r \quad (24)$$

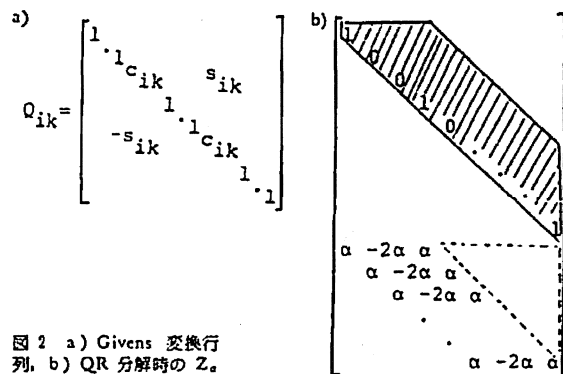


図2 a) Givens 変換行列, b) QR 分解時の $Z_\alpha$

$Q_{ik}$ を $Z_\alpha$ の左から作用させると、 $k$ 行と $i$ 行のみが、

$$\begin{aligned} z_{kj} &= c_{ik} z_{kj} + s_{ik} z_{ij}, \quad k \leq j \leq n, \\ z_{ij} &= -s_{ik} z_{kj} + c_{ik} z_{ij}, \end{aligned} \quad (25)$$

と変形される。 $Z_\alpha$ の構造から fill-in の起こる場所があらかじめ分り、なるべく fill-in を押さえるようにピボットを選択することができる(このことがDを図1-c, d)とした理由である)。

fill-in を押えるために $\|y - E f\|$ を $\|E^* y - E^* E f\|$ と変形して考える。 $E^* E$ は対角要素が0か1( $E^* y$ の要素にデータがあるかないかに対応する)で、非対角要素がすべて0となるような行列になる。

そして、Dの要素を上から一行ごとに生成し(Dを記憶しておく必要がないことに注意)、Givens 変換によりその非零要素を消去する。このときDの非零要素の長さ(帯幅)以上には fill-in は起こらない。その理由は $Z_\alpha$ の非零要素が $E^* E$ の対角成分にしかなく、Dを図1-c, d)のようにとっているからである。したがって、QR 分解に必要な記憶容量は最終形がRとなる図2-b)の斜線の部分のみである。なお、Householder 変換法を用いた場合は点線内の部分すべてに fill-in が起こる。

$\hat{f}$ を求めるアルゴリズムは次のようになる。

アルゴリズム 1.

ステップ 1:  $\alpha$ を定める。

ステップ 2:  $z_\alpha$ を Givens 変換で QR と分解する。

ステップ 3: ABIC 最小ならステップ5へ

ステップ 4:  $\alpha$ を変更してステップ2へもどる。

ステップ 5:  $R_j f = b_j$  を計算して  $f$  を求める。

ただし、実際の計算上では ABIC ( $\alpha$ ) は非線形で凸性の保証がないので (実験的には、“ほとんど” 凸である) 大域的にいくつかの点を選んで ABIC ( $\alpha$ ) を計算し、ある程度区間を狭めてから黄金分割による区間縮小法を用いる。ただし、 $\alpha$  は必ずしも厳密に求める必要がないことに注意。

#### 4. 離散スプライン

この節では、 $f$  の持つ性質について述べる。簡単のためにまず一次元問題について考える。(この節ではデータのある fit-point のことを data-point と書くことにする)。(20) 式は

$$\min_f (\| \bar{y} - Ef \| + \alpha^2 \| Df \|^2) \quad (26)$$

と書ける。したがって、これは次のような連立方程式を解くことと同値である。

$$\alpha^2 D^* D f = E^* E f - E^* \bar{y} \quad (27)$$

(27) の各方程式は  $x_j$  が data-point の場合と fit-point の場合の 2 種類の方程式に分けることができる (ただし、両端 2 点ずつは除く)。

$$\textcircled{1} \text{ (data-point の場合) } \Delta^4 f_j = (f_j - \bar{y}_j) / \alpha \quad (28)$$

$$\textcircled{2} \text{ (fit-point の場合) } \Delta^4 f_j = 0 \quad (29)$$

$$\text{ここで } \Delta^4 f_j = f_{j-2} - 4f_{j-1} + 6f_j - 4f_{j+1} + f_{j+2} \quad (30)$$

②の場合、 $f_j$  において  $f$  の 4 回差分が零ということ、 $f_{j-2}, f_{j-1}, f_j, f_{j+1}, f_{j+2}$  が同じ三次式に乗っていることを示している。また、 $\Delta^4 f_j = 0, \Delta^4 f_{j+1} = 0$  であれば、2 つの三次式は 4 点を共有しその 4 点で三次式が決定するので、 $f_{j-2}$  から  $f_{j+2}$  までが同一の三次式上の点となる。つまり、 $f$  は隣りあう data-point 間とそれらの data-point の外側の fit-point までの間がひとつの三次式にのっていることになる (図 3 の  $f_{j-1}$  から  $f_{j+4}$  までの点)。

data-point  $f_{j-1}$  と  $f_j$  間と data-point  $f_j$  と  $f_{j+1}$  間の 2 つの三次式は  $f_{j-1}, f_j, f_{j+1}$  で交わっているが、その微係数は一致していない (ただし、 $f_j$  でのみ二階微係数が一致している)。data-point の両端にある fit-point が data-point に近ければ近い程  $f$  はなめらかな関数を表現しうることになる。これらのことからこ

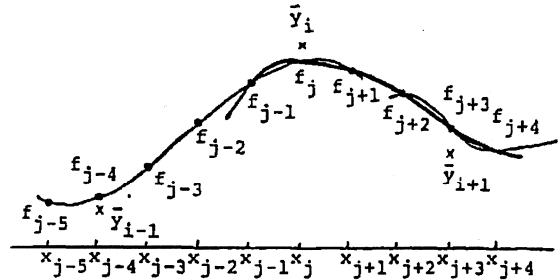


図 3 離散スプラインの概念図

のモデル  $f$  を三次の「離散スプライン」と呼ぶことにする。 $f$  は連続型モデル (三次スプライン関数) をサンプリングしたものではない。(28~30) 式は  $\alpha$  の値にかかわらず成り立つことに注意。

fit-point を増やした場合は、データ間で新しい三次式が生成される。これは fit-point を増やす前の三次式とは一致しない。

離散スプラインを用いる時の注意として、両端の影響をできるだけ少なく、かつデータの  $f$  に与える影響の仕方をすべて同等にするために最左端の data-point の左に 2 点の fit-point、最右端の data-point の右に 2 点の fit-point をつねに取るのがよい。また fit-point 間の内挿が要求される場合には data-point 間を三次式とすることが自然であろう。

離散スプラインが三次式になるのは  $D$  として二階差分を用いているからである。 $D$  として  $n$  階差分を用いれば離散スプラインは  $2n-1$  次式となることは (27) 式より容易にわかる。

二次元モデルの場合、一次元モデルの (30) 式に対応する式は、

$$\begin{aligned} \Delta^4 f_{1,j} = & f_{1,j-2} + 2f_{1,j-1} - 8f_{1,j} + 2f_{1,j+1} \\ & + f_{1,j+2} - 8f_{1,j+1} + 20f_{1,j} - 8f_{1,j-1} + f_{1,j-2} \\ & + 2f_{1,j+1} - 8f_{1,j} + 2f_{1,j-1} + f_{1,j+2} \end{aligned} \quad (31)$$

となる。

二次元モデルでは一次元モデルのときのように隣りの点を中心としたものと同じ曲面にはならない。

また、一次元モデルのとき両端に 2 ずつ fit-point をとったように二次元モデルでは二重の fit-point の枠をとるのがよい。

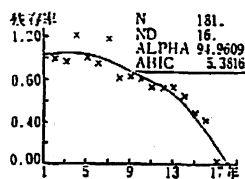


図4 離散スプラインによるあてはめ（ビール箱の市場残存率）。N は fit-point 数, ND は data-point 数

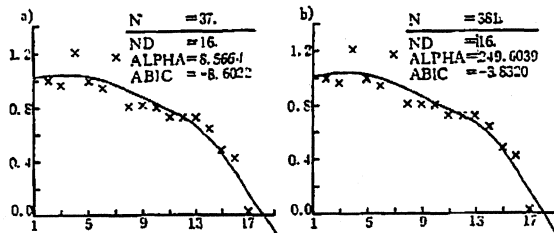


図5 fit-point 数nによる $\alpha$ の変化。a) nを小さくした場合, b) nを大きくした場合

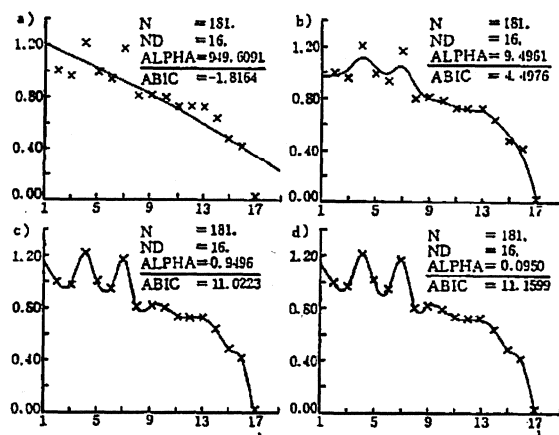


図6  $\alpha$ を変えた時の離散スプラインの挙動

## 5. 適用例

### 1) 一次元問題

一次元の問題にアルゴリズム1を適用した例をみる。

図4はビール箱の市場残存率を示す統計データに離散スプラインをあてはめた例である。fit-pointは、データ数の10倍取ってある。

図5-a)は同じデータに対してfit-pointの数nを変化させてみたものである。ABICを用いるとfit-pointの数nが多くなるに従って $\alpha$ の値を大きくしてfのなめらかさを調整していることが観察される。

図5-b)は $\alpha$ の値を人為的に定めて、 $\alpha$ による離散スプラインの挙動を調べたものである。(26)式からもわ

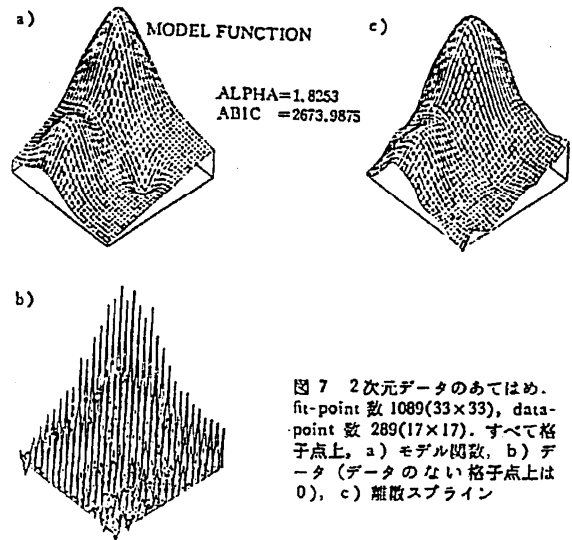


図7 2次元データのあてはめ。fit-point 数1089(33×33), data-point 数289(17×17)。すべて格子点上。a) モデル関数, b) データ(データの無い格子点上は0), c) 離散スプライン

かるように、 $\alpha$ が最適値 $\hat{\alpha}$ より大きいと近似関数fのなめらかさを強くするのであるからfは直線に近づく。 $\alpha$ を十分大きくとれば、線形回帰に近づく(図6-a))。

逆に $\alpha$ の値を小さくしていくと、この場合は先程とは逆にfのなめらかさを弱くして離散スプラインはデータの動きに追随するようになってくる。十分に $\alpha$ の値を小さくしていくとfはデータ補間に近づいていく(図6-3b), c), d))。

### 2) 二次元問題

一般に、一次元問題の場合とちがって二次元の問題に対してのあてはめや補間は近似関数に取りにくく難しい[4]。しかし、本稿のモデルでは一次元問題の場合とおなじアルゴリズムを用いることができる。図7に結果を示す。図7-a)は2つの山と1つの谷を持っているモデル関数である。このモデル関数を1089(33×33)の格子点上のfit-pointであてはめてみる。図7-b)はモデル関数に正規誤差を加えたもので、これをデータとする。データは289(17×17)の格子点上にある。このデータをあてはめた結果が図7-c)である。

次に $\alpha$ の値を変えてABICの値を比較した結果を図8に示す。fit-pointは前図と同じ格子点上ですべてのfit-point上に誤差を乗せたデータがある。一次元の場合と同様に $\alpha$ が小さいと振動が激しく、 $\alpha$ が大きくなると山や谷が消えて面が平になってしまう。ABIC最小の時が(5つの中で)、一番モデルに似て

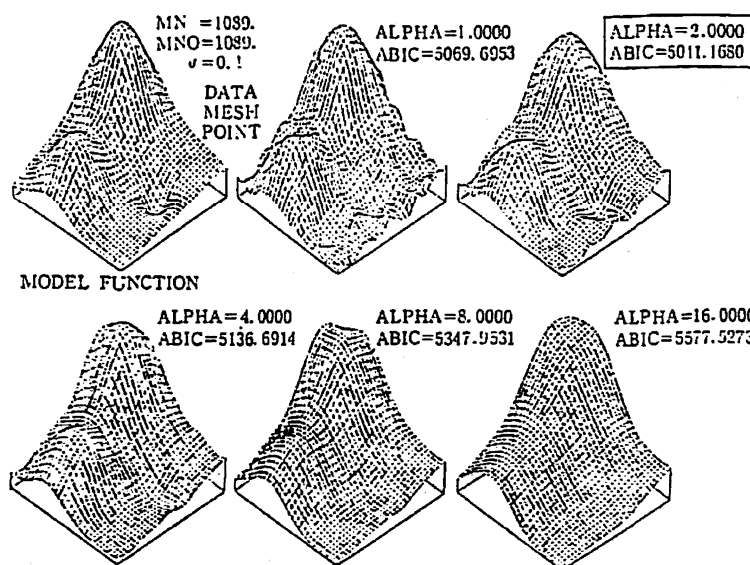


図 8  $\alpha$  を変えた時の離散スプラインの挙動

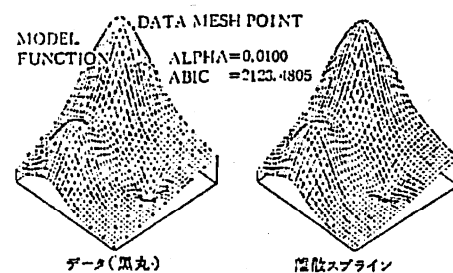


図 9 データ補間としての離散スプライン (データ数 81(9x9))

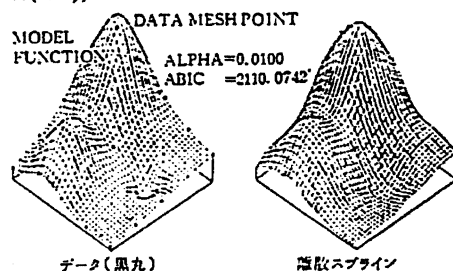


図 10 データ補間としての離散スプライン (データ数 25(5x5))

いることがわかる。  $\alpha$  を小さくするとデータ補間に近づくことを一次元問題でみたが、これを利用して二次元データ補間に離散スプラインを用いることを試みる。図 9、図 10 はデータに誤差をのせずに補間した例である。データは格子点 81(9x9) の場合と 25(5x5) の場合を示した。

## 6. 二相問題のあてはめ (一次元問題)

あらかじめ「構造」がある点  $x_p$  で 2 つの相に分けられることがわっている問題について考察する (二相問題)。ここでは 2 つの問題点について考える。1 つは分岐点  $x_p$  の両側をそれぞれ別の相としてあてはめること。もうひとつは  $x_p$  の位置を推定することである。

まず、前者の問題について考える。求める区間を  $n$  分割し、  $x_r$  で二相に分かれているとモデリングする離散スプラインは  $f$  のなめらかさを事前分布として扱っているが、この事前分布の仮定を次のように変更する。  $D$  を 3 つの部分 (第 1 行から第  $r-1$  行まで、第  $r$  行、第  $r-1$  行から第  $n-2$  行まで) に分離し、それぞれに  $\alpha, \beta, \gamma$  という重みを超パラメータとして対応させる。3 節での議論を繰り返すと (26) 式に対応する式は、

$$\min_f (\|y - Ef\|^2 + \alpha \|D_1 f\|^2 + \gamma \|D_2 f\|^2 + \beta \|D_3 f\|^2) \quad (32)$$

となる。このときの ABIC は、

$$\begin{aligned} \text{ABIC}(\alpha, \gamma, \beta, r) = & (N-2) \log (\|b - Z_{\alpha\beta\gamma} f_*\|^2) \\ & + \log |\det(Z_{\alpha\beta\gamma} Z_{\alpha\beta\gamma})| \\ & - 2(r-1) \log \alpha - 2 \log \gamma \\ & - 2(n-3+r) \log \beta + C \end{aligned} \quad (33)$$

となり、アルゴリズム 1 を用いてあてはめを行なうことができる。

$x_p$  の位置の推定の問題に対しては  $f$  をパラメータとして扱い、  $r$  を 3 から  $n-3$  まで動かし (端の近くを選ぶときは注意が必要)、ABIC を比較して最適な  $r$  を求める。このとき  $x_r$  が分岐点として選ばれることになる。よってこの方法では  $x_p$  を厳密に求めるのではなく  $x_p$  に近い点  $x_r$  を求めることになる。

あてはめ例をみよう。例題はポリアクリルニトリル中の  $m$ -ニトロアニリンの拡散係数  $d$  の Arrhenius プロット図である。この問題はあらかじめ相が変わることがわかっている。まず、一相問題として離散スプラインであてはめたものが図 11-a) である。図 11-b) は二相問題として  $r$  を変えて各  $r$  に対して ABIC 最小の値を図示したものである。最適の  $r$  の時の二相あ

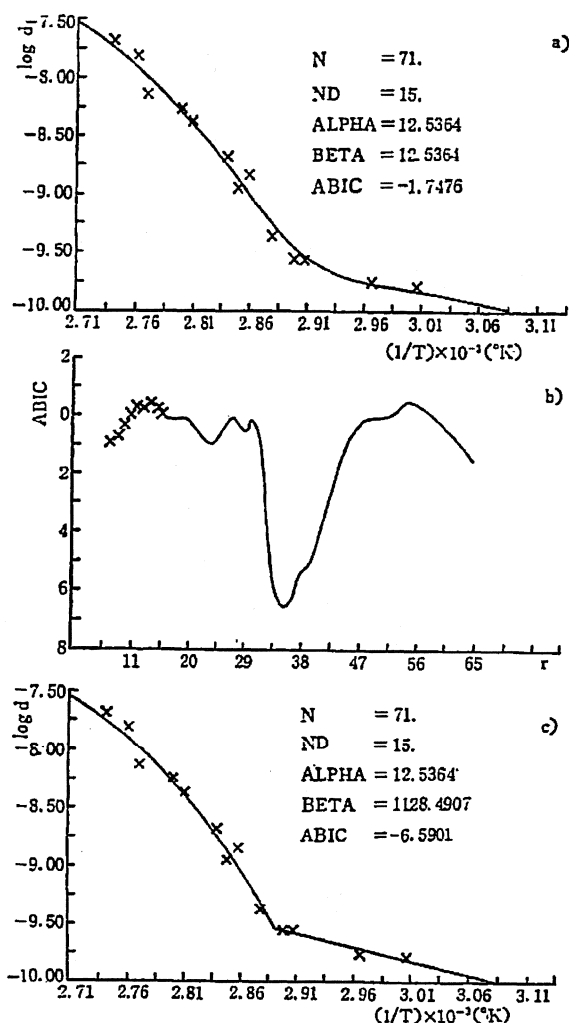


図 11 a) 離散スプラインによるあてはめ (問題を一相として), b)  $r$  を変えた場合の  $\text{ABIC}$  変化, c) 最適な  $r$  で二相に分けた場合のあてはめ

てはめの結果を図 11-c) に示す。

スプライン関数で節点の多重度を上げて不連続点を表現できるように、本稿の離散スプラインでも、第  $r$  行、第  $r+1$  行を 2 つ続けて  $D_2$  として (32) 式の第 2 項とすれば、その点で不連続な関数を表現できる。

離散スプラインの利点は相の分岐点を  $\text{ABIC}$  によって、評価できることである。また相に分けられるかどうかをも  $\text{ABIC}$  を比較することに行なうことができそうである。

本稿は田中・田辺<sup>[17]</sup>に基づくものであり、原論文では、非負性、単調性、凸性などの構造情報を数理計画問題に定式化して  $f$  のより良い推定を得る方法など

についてより詳しく述べているので参照いただきたい。

#### 参考文献

- [1] Akaike, H.: Likelihood and Bayes procedure, In Bayesian Statistics, J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds, University Press, Valencia, Spain, 143-166 (1980).
- [2] 赤池弘次: 統計的推論のパラダイムの交差について, 統計数理研究所集報, 27, No. 1, 5-12 (1980).
- [3] 赤池弘次: モデルによってデータを測る, 数理科学, 218, 7-10 (1981).
- [4] Franke, R.: A critical comparison of some methods for interpolation of scattered data, Naval Postgraduate School, Monterey, California, NPS-53-79-003 (1979).
- [5] Ishiguro, M., Akaike, H.: Trend estimation with missing observations, Ann. Inst. Statist. Math., 32, Part B, 481-488 (1980).
- [6] 石黒貞木夫: ベイズ型季節調整モデル, 数理科学, 218, 57-61 (1981).
- [7] 石黒貞木夫, 荒畑恵美子: ベイズ型スプライン回帰, 統計数理研究所集報, 30, No. 1 (1982).
- [8] 柏木直久: 臨場試験に於ける地力の推定, 統計数理研究所集報, 30, No. 1, 1-10 (1982).
- [9] 北川源四郎: 異常値解析ベイズモデル, 数理科学, No. 218, 62-66 (1981).
- [10] Kitagawa, G., Akaike, H.: A Quasi Bayesian Approach to outlier detection, Ann. Inst. Statist. Math., 34, Part B, 389-398 (1982).
- [11] 中村 隆: ベイズ型コウホート・モデル-標準コウホート変への適用-, 統計数理研究所集報, 29, No. 2, 77-98 (1982).
- [12] 田辺國士: 統計データの誤差の処理, bit 臨時増刊「数値計算における誤差」, 113-125 (1975).
- [13] 田辺國士: 不適切問題の統計的および数値的取り扱いについて, 日本数学会, 応用数学分科会, 講演予稿集 (1975).
- [14] 田辺國士: 不適切問題への統計的アプローチ, 数理科学, No. 153, 1-5 (1976).
- [15] Tanabe, K.: Statistical regularization of a noisy ill-conditioned system of linear equations by Akaike's information criterion, Computation and Analysis, 6, No. 4, 2-25 (1975).
- [16] 田中輝雄: 最小二乗問題における Givens 法と Householder 法について, 統計数理研究所集報, 30, No. 2, 掲載予定 (1983).
- [17] 田中輝雄・田辺國士: ベイズの方法によるデータのあてはめ, 京大数理研究所研究録 (1983).

# Statistical procedure for deconvoluting experimental data

Yuzuru Hiragi and Hiroshi Urakawa

*Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto-fu, 611 Japan*

Kunio Tanabe

*Institute of Statistical Mathematics, Minami-Azabu, Minatoku, Tokyo 106, Japan*

(Received 26 November 1984; accepted for publication 12 February 1985)

A general procedure of the deconvolution of noisy experimental data is proposed. The deconvoluted solution is represented in terms of a cubic spline function  $f(x)$  with variable knots. The value of parameters and the location of knot points in the spline are determined by solving the nonlinear least-squares problem of fitting  $g(x) = \int \kappa(\tau)f[s(\tau, x)]d\tau$  to the data, where experimental data  $g(x)$  represent the convolution of  $f[s(\tau, x)]$  with an apparatus weighting function  $\kappa(\tau)$  and  $s(\tau, x)$  is a function related to the configuration of the experimental system. The profile of the fitted solution varies widely with different choices of the number and position of knots in the spline. To obtain a good solution, the optimum number of knots is determined based on the data by applying minimum Akaike's Information Criterion procedure. Practical application of the variable-knot spline is generally considered to be rather intricate. However, by taking advantage of an advanced technique of a quasi-Newton method, the nonlinear least-squares problem becomes tractable, and excellent smoothing in the deconvoluted solution can be achieved.

## I. INTRODUCTION

For many physical and chemical experiments, data are often available exclusively as the convolution of instrumental functions prescribed by the experimental system with an ideal curve, which requires analysis. For example: absorption spectra are given as the convolution of an ideal curve with slit width<sup>1</sup>; x-ray<sup>2</sup> and neutron small-angle scattering<sup>3</sup> curves with profile of the primary beam and wavelength; fluorescence decay curves<sup>4-6</sup> with the time structure of the excitation; elution patterns of gel chromatography with the profile of molecular weight distribution; and heat capacity measurement curves with the mutual convolution of the multistep transitions.<sup>7-9</sup>

Many methods have been proposed for the deconvolution of data which arise in various branches of scientific research. Mathematically, the problem reduces to one of solving the Fredholm-type integral equation of the first kind,

$$g(x) = \int \kappa(\tau)f[s(\tau, x)]d\tau, \quad (1)$$

where  $g(x)$  is the measured spectrum,  $\kappa(\tau) = \kappa(\tau_1, \tau_2, \dots, \tau_n)$  is the instrumental weighting function,  $f(x)$  is the true curve, and the function  $s(\tau, x)$  reflects, for example, the geometrical configuration of the experimental system. Among the existing methods, that of iteration<sup>1,10</sup> and of the Fourier transform<sup>11</sup> are the most popular. However, direct solution of Eq. (1) often leads to oscillation in the solution curve  $f(x)$ , which does not reflect true structure, since the data contain errors (noise). The matter then comes under the category of ill-posed problems.

Instead of solving Eq. (1) directly, least-squares method were developed for the problem by using a spline as the model function.<sup>3,12</sup> However, there is still difficulty in choosing an appropriate number of free parameters in the model. If too few parameters are allowed in the model, some of the information contained in the data will be lost. Too many parameters will cause oscillation in the estimated solution

$f(x)$ . To avoid the latter situation, a subsidiary constraint is usually introduced and the corresponding Lagrange multiplier  $\lambda$  is used to control the smoothness of  $f(x)$ . However, the choice of the value  $\lambda$  is, in most cases, empirical and often subjective. Thus, the criterion for objective determination of an appropriate smoothness in the solution has been sought. In this paper, we propose an automatic method for determining smoothness in the deconvoluted solution by applying AIC (Akaike's Information Criterion) to choose an optimum number of knot points in the cubic spline approximation with variable knots. An earlier attempt was made on the same line<sup>23</sup> of research. The procedure of foolproof in the sense that smoothness in the solution is determined objectively based only on the data without requiring any judgement by investigators. Fitting capability of the cubic spline is much enhanced by introducing the variable (free) knots. Although the resulting nonlinear least-squares problem requires a rather expensive computing effort, we found it worth the cost.

## II. MATHEMATICAL BACKGROUND

### A. Least squares method as a maximum likelihood method for the data containing the errors with normal distribution

For a set of  $N$  data points  $(y, x) = (y_i, x_i)$ ,  $i = 1, 2, \dots, N$ , the  $i$ th data  $y_i(x_i)$  consists of true value  $y_i^0$  and error  $e_i$ ,

$$y_i = y_i^0 + e_i, \quad (2)$$

where  $e_i$  is independently as well as identically distributed with mean 0 and variance  $\sigma_i^2$ , and  $y_i^0$  and  $\sigma_i$  are known. Generally, the distribution of experimental errors  $e_i$ 's are expressed or well approximated in terms of a normal (gaussian) distribution. Hence, the density function of  $e_i$  is given by

$$(2\pi\sigma_i^2)^{-1/2} \exp(-e_i^2/2\sigma_i^2), \quad (3)$$

and the likelihood function of  $y_i^0$  in this case is

$$L(y_i^0) = (2\pi\sigma_i^2)^{-1/2} \exp[-(\bar{y}_i - y_i^0)^2/2\sigma_i^2], \quad (4)$$

where  $\bar{y}_i$  is the observation of  $y_i$ . The likelihood function for the vector  $y^0 = (y_1^0, \dots, y_N^0)'$  is given by

$$L(\bar{y}|y^0) = (2\pi)^{-N/2} \left( \prod_{i=1}^N \sigma_i^2 \right)^{-1/2} \times \exp \left[ -\frac{1}{2} \sum_{i=1}^N (\bar{y}_i - y_i^0)^2 / \sigma_i^2 \right], \quad (5)$$

where  $\bar{y} = (\bar{y}_1, \dots, \bar{y}_N)'$  and  $y'$  denotes the transpose of  $y$ . The maximum likelihood estimate  $\hat{y}$  of  $y^0$  is given by maximizing the likelihood function, Eq. (5). This is achieved by minimizing  $-\log L(\bar{y}|y)$  with respect to  $y$ , hence, by solving the least squares problem,

$$vv' = \sum_{i=1}^N (\bar{y}_i - y_i)^2 / \sigma_i^2 \rightarrow \text{minimum}. \quad (6)$$

The least-squares method is equivalent to the maximum likelihood method under the assumption that errors are normally distributed

## B. Expression of the solution by a model function

Let us now express the estimated value  $\hat{y}_i$  by a model function  $\hat{g}(x)$  in the following way,

$$\hat{y}_i = \hat{g}(x_i). \quad (7)$$

$\hat{g}(x)$  can be conveniently represented by the integral transformation of  $\hat{f}(x)$  [see Eq. (1)], where  $\hat{f}(x)$  is a model function. Let  $\hat{f}(x)$  be the linear combination of  $M$  constituent bases  $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_M)'$  with  $a_j$  as coefficient,

$$\hat{f}(x) = \sum_{j=1}^M a_j \varphi_j(x). \quad (8)$$

Substituting Eq. (8) into (1) we have,

$$\hat{g}(x) = \sum_{j=1}^M a_j \psi_j(x), \quad (9)$$

where

$$\psi_j(x) = \int \kappa(\tau) \varphi_j(s(\tau, x)) d\tau. \quad (10)$$

Generally  $\varphi_j$  is expressed in terms of a function of  $s(\tau, x)$  which depends on the experimental condition [see, e.g., Eq. (17)]. Since the integral transformation (1) is linear, we determine  $a_j$  by solving the linear least-squares problem,

$$vv' = \sum_{i=1}^N \frac{1}{\sigma_i^2} [\bar{y}_i - \hat{g}(x_i)]^2 = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[ \bar{y}_i - \sum_{j=1}^M a_j \psi_j(x_i) \right]^2 \rightarrow \text{minimum}. \quad (11)$$

Although we can use any set of functions as  $\varphi_j(x)$  (for example, Fourier series, Tchebycheff series, etc.), the  $B$  spline<sup>13,14</sup>

$$\varphi_j(x) = (t_{j+k} - t_j) \sum_{r=j}^{j+k} \frac{(t_r - x)_+^{k-1}}{(t_r - t_j) \cdot (t_r - t_{r-1}) \cdots (t_r - t_{j+k})}, \quad (14)$$

where  $t_r$  is a knot and

$$(t_r - x)_+^{k-1} = \begin{cases} (t_r - x)^{k-1} & x < t_r \\ 0 & x > t_r \end{cases}$$

seems to be the most appropriate for curve fitting. From Eq. (1), the normal equation for  $\alpha = (\alpha_1, \dots, \alpha_M)'$  is derived as

$$G = Fa, \quad (12)$$

where

$$G = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_M \end{bmatrix}, F = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1M} \\ F_{21} & F_{22} & \cdots & F_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ F_{M1} & F_{M2} & \cdots & F_{MM} \end{bmatrix}$$

with matrix elements,

$$G_j = \sum_{i=1}^N \bar{y}_i \psi_j(x_i) / \sigma_i^2$$

and

$$F_{jj'} = \sum_{i=1}^N \psi_j(x_i) \psi_{j'}(x_i) / \sigma_i^2.$$

Consequently,  $a_j$  is given by

$$a = F^{-1}G. \quad (13a)$$

$F^{-1}$  is the inverse matrix of  $F$ , i.e.,

$$a = (\psi' W \psi)^{-1} \psi' W \bar{y}, \quad (13b)$$

where

$$\bar{y} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_N \end{bmatrix}, \psi = \begin{bmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1M} \\ \psi_{21} & \psi_{22} & \cdots & \psi_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{N1} & \psi_{N2} & \cdots & \psi_{NM} \end{bmatrix} \text{ and } \psi_{ij} = \psi_j(x_i)$$

and the weight matrix is given as

$$W = \begin{bmatrix} \sigma_1^{-2} & & & \\ & \sigma_2^{-2} & & \\ & & \ddots & \\ & & & \sigma_N^{-2} \end{bmatrix}.$$

In practice, the value of  $a$  is computed by the Householder transformation instead of solving Eq. (12) numerically. The minimization of Eq. (11) with respect to locations of knot points will be discussed in Appendix A.

## C. Spline with variable knots

The  $(k-1)$ th power  $B$  spline represents a normalized local base  $\varphi_j$  for the spline function  $\hat{f}(x)$  in the present method:

$k=4$  in Eq. (14) corresponds to a cubic spline.

Although knots are fixed in the usual spline applications, the benefit of the spline function is best given by allow-

ing the knots to move to the positions which yield a nicer spline approximation to the data.<sup>13,14</sup> This means that the vector,  $\mathbf{t} = (t_1, t_2, \dots, t_M)'$ , as well as  $a$ , are to be considered as a variable in the least-squares procedure. Thus, we will denote, hereafter,  $\varphi_j(x)$  and  $f(x)$  as  $\varphi_j(x) = \varphi_j(x, \mathbf{t})$  and  $\hat{f}(x) = \hat{f}(x, \mathbf{a}, \mathbf{t})$ , respectively. As is easily seen from Eq. (14),  $t_i$ 's behave nonlinearly in  $\hat{g}(x)$ . Hence, a nonlinear least-squares problem should be solved. For this purpose, we use the procedure (see Appendix A) based on a modification<sup>15</sup> of the quasi-Newton iterative method due to Biggs,<sup>16</sup> which was implemented in a computer program NOLLS1.<sup>17</sup>

#### D. Determination of model by minimum Akaike's information criterion estimation procedure

An ultimate purpose of our least-squares procedure is to find a model function  $\hat{f}(x)$ , which derives the full information from the observed data and has least number of parameters. This is achieved firstly by proper selection of  $\varphi_j(x)$ , for which we have adopted a cubic  $B$  spline, secondly by choosing optimal positions for knots as carried out by our nonlinear least-squares algorithm, and thirdly by determining an appropriate number of parameters (knots) as realized by minimizing the statistics,  $AIC^{18-22}$  (see Appendix B),

$AIC = -2 \ln$  maximized likelihood of a model with respect to given data + Degree of freedom of the model. (15)

From the discussion given in Sec. II A, we have

$$W_l(t) = \begin{cases} 1 & \text{for } -A < t < A \\ \pm t / (B - A) + B / (B - A) & \text{for } -B < t < -A \text{ and } B > t > A \\ 0 & \text{for } t < -B \text{ and } t > B \end{cases}$$

and

$$W_w(u) = (2\pi\sigma)^{-1/2} \exp(-u^2\sigma^2/2),$$

with  $A$  and  $B$  being positive real numbers. For a test function of the scattering curve (true curve), we adopted the Lorentzian type as  $f_0(x)$ :

$$f_0(x) = \sum_{i=1}^3 \frac{d_i}{a_i(x - b_i)^2 + c_i}, \quad (18)$$

where  $(\alpha_i, b_i, c_i, d_i) = (1.5, 0.0, 0.25, 2.0)$  for  $i = 1$ ,  $= (4.0, 1.6, 0.25, 0.4)$  for  $i = 2$  and  $= (6.0, 2.4, 0.4, 0.4)$  for  $i = 3$ . "Noiseless data" or true value  $y_0(x) = [y_0(x_i)]$  were obtained by the convolution of  $f_0(x)$  with  $\kappa(t, u)$  in terms of Eq. (17). "Simulated experimental data"  $y(x)$  was generated by the formula,  $y(x_i) = y_0(x_i) + e_i(x_i)$ , where  $e_i$  is normally independently and identically distributed as expressed in terms of  $N\{0; [y_0(0x_i)]^{1/2}\}$ .

In Fig. 1(a), the sum of squared residuals and  $AIC$ 's are plotted against the number of knots in the spline. Whereas the sum of squared residuals is a monotone decreasing function of the number of knots, the  $AIC$  value behaves differently. The deconvoluted solution with the optimum number of knots (9 knots) obtained by MAICE is shown in Fig. 1(c), together with one with fewer knots [Fig. 1(b)] and another with more knots [Fig. 1(d)]. The recovered curve resembles

$$AIC = N \ln(\mathbf{v}\mathbf{v}') + 2M. \quad (16)$$

As  $AIC$  represents the amount of the deviation of the model from the true curve, we choose the one with the smallest  $AIC$  value among competitive models with varying degrees of freedom<sup>18,21-23</sup>. By this method which is called MAICE (minimum  $AIC$  estimation), we are able to determine the numbers of  $\varphi_j$ ,  $M$ , and positions of knots without resorting to subjective judgement. The generalized cross-validation method<sup>24</sup> could also be used for this purpose. We adopt MAICE here for its generality.

To find the smallest  $AIC$  value, there are two phases of minimization. One is to solve the linear least-squares problem with respect to  $\mathbf{a}$  under the condition that  $t_i$ 's are fixed, and the other is to minimize  $\mathbf{v}\mathbf{v}'$  with respect to the location of the knots points.

### III. APPLICATION TO DESMEARING AND DISCUSSION

As a numerical example of the deconvolution, we applied our method to the following equation,

$$g(x) = \int \kappa(t, u) f\{[(x - u)^2 + t^2]^{1/2}\} du dt. \quad (17)$$

Eq. (17) simulates the simplified case of a small-angle scattering curve by a line-shaped primary x-ray beam profile,<sup>25,26</sup> where the scattering curve  $f(x)$  of pinhole collimation system is "smeared" by the beam profile  $\kappa(t, u)$ . Here  $\kappa(t, u) = W_l(t)W_w(u)$ ,  $W_l$  has a trapezoid form,

the original curve very well. Fig. 1(e) shows the profile of least-squares spline fit with fixed knot and Fig. 1(f) shows that with variable knot. In contrast to the fixed knot spline, the advantage of the variable knot is well demonstrated in this example.

Figure 2 shows the optimum profile obtained from the data with magnified variance (hence errors were increased) by the factor of  $10^{1/2}$ . In this case, the MAICE procedure chooses a spline with fewer knots as the best model. The recovered curve is somewhat deteriorated as compared with the case demonstrated in Fig. 1(c). This fact is consistent with our intuition that the information content is more corrupted by larger errors and the resultant profile of the solution should be less precise. This behavior is typical of the MAICE procedure.

Mathematically, the present procedure aims to smooth the experimental data which are considered to be the integral transforms of a certain spectrum. Thus, its application is not restricted to deconvolution but is extended to other integral transformations such as Fourier, Hankel, Bessel, etc. A similar application was made to the Hankel transform of a small-angle scattering intensity curve by Glatter,<sup>27</sup> but in this method the stabilization of the solution is made empirically. Lastly, we would like to point out that the application to the Fourier transformation case, where sine and cosine functions are used instead of spline function, would be interesting



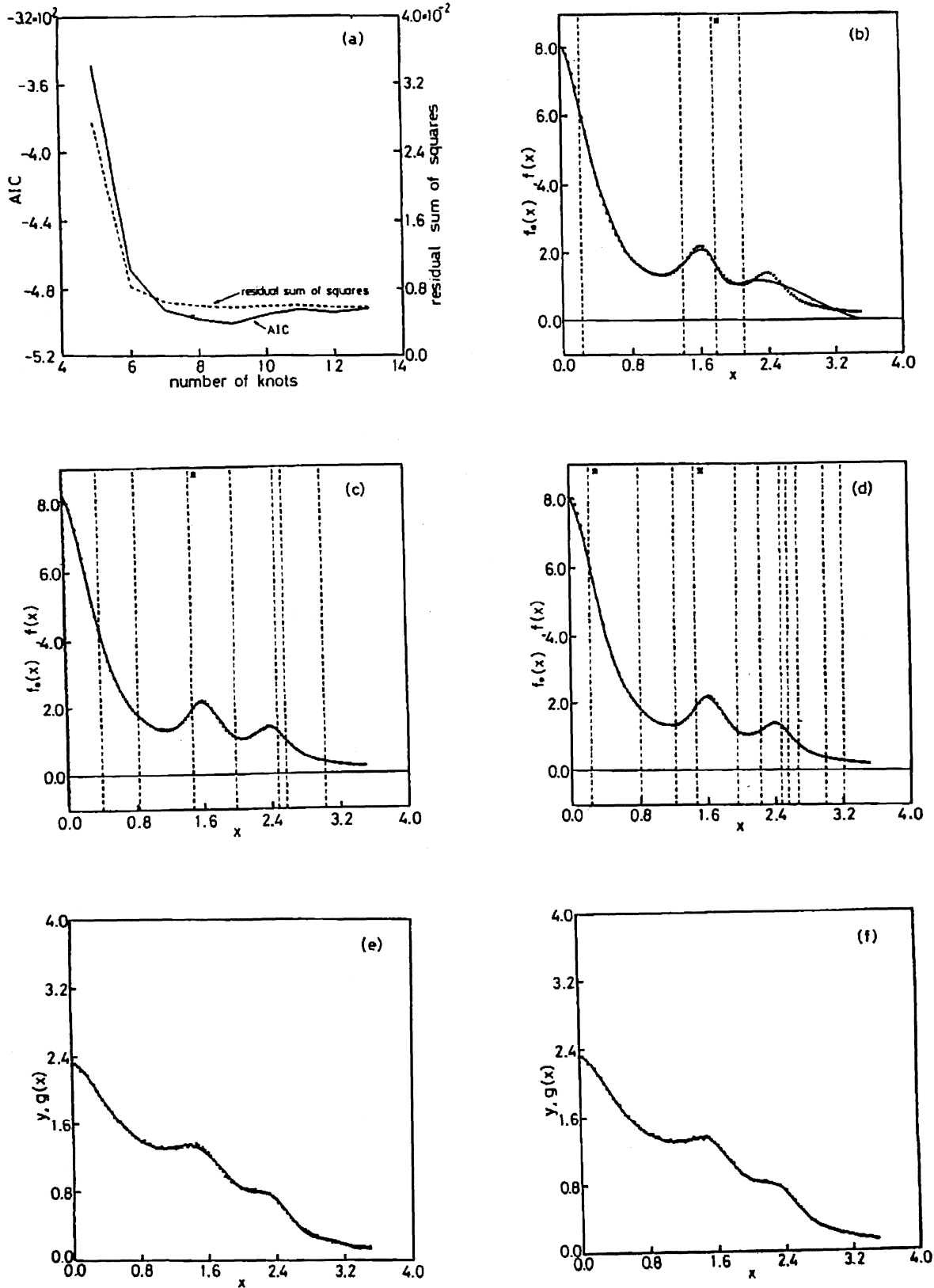


FIG. 1. (a) Plot of optimized value of AIC (solid line) and residual sum of squares (broken line) against number of knots. Results of deconvolution with different number of knots; (b) with 5 knots, (c) with 9 knots (AIC minimum), (d) with 13 knots. Results of deconvolution is indicated with — and original  $f_0(x)$  curve expressed with Eq. (18) with ·····. Vertical lines indicate the position of knots and the line marked by \* is a multiple knot points. Fixed (linear) (e) and optimized (f) least squares fit with 9 knots in data space; for both (e) and (f), ····· is "experimental data" and — fitted  $\hat{g}(x)$  curve.

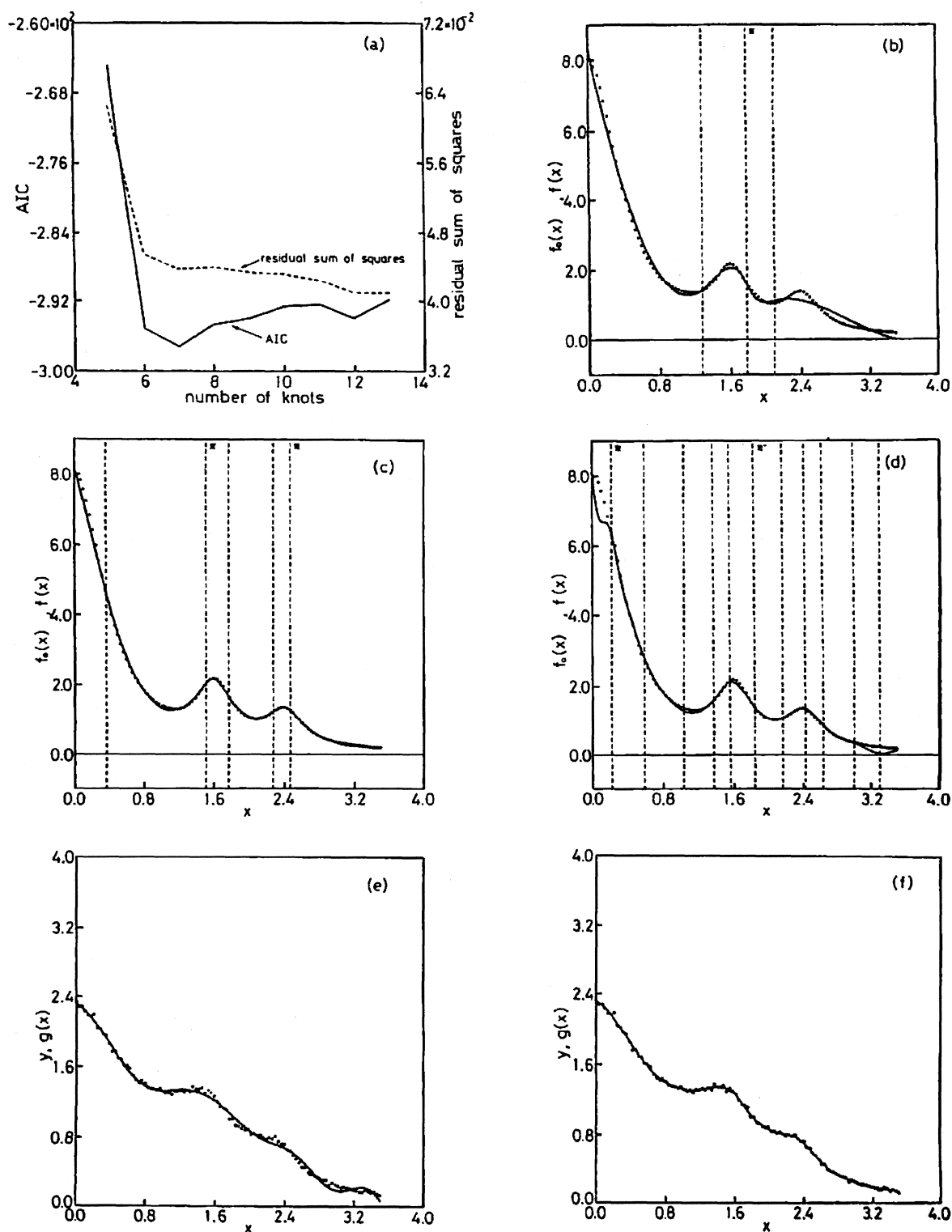


FIG. 2. Similar figures to Fig. 1 but variances (errors) of the "data" are increased to  $10^{1/2}$  times. Consequently, optimum number of knots decreased to 7.

since it is used frequently and the resulting linear least-squares problem is easy to solve. The application of the MAICE procedure to this case is straightforward and the proper number of series functions is easily determined.

#### ACKNOWLEDGMENT

We are grateful to Dr. H. Fujii and Dr. K. Ichida for their discussions during the early stage of this work. One of

us (H.U.) would like to thank Dr. R. Kitamaru and Dr. K. Kaji for their continual guidance. Thanks are also due to Dr. T. Ooi for his interest and Dr. K. Kajiwaru for reading the manuscript.

## APPENDIX A: QUASI-NEWTON METHOD FOR NONLINEAR LEAST-SQUARES PROBLEM

For each choice of  $M$ , we are required to minimize the residual sum of squares,

$$V(\mathbf{a}, \mathbf{t}) = \frac{1}{2} \mathbf{v} \mathbf{v}' = \frac{1}{2} \sum_{i=1}^N v_i^2 = \frac{1}{2} \sum_{i=1}^N \left( \frac{\bar{y}_i - g_i}{\sigma_i} \right)^2, \quad (\text{A1})$$

with respect to the variable (position of knots),  $\mathbf{t} = (t_1, t_2, \dots, t_M)'$ , as well as the variables  $\mathbf{a}$ , where  $g_i = g(x_i, \mathbf{a}, \mathbf{t})$  is a transformed cubic  $B$ -spline function evaluated at  $x_i$ . Since  $v_i$ 's are a linear function of  $\mathbf{a}$ ,  $V$  is easily minimized with respect to  $\mathbf{a}$  when  $\mathbf{t}$  is fixed. Let

$$W(\mathbf{t}) = \min_{\mathbf{a}} V(\mathbf{a}, \mathbf{t}) = V(\hat{\mathbf{a}}(\mathbf{t}), \mathbf{t}) \quad (\text{A2})$$

and let

$$\mathbf{w}(\mathbf{t}) = [w_1(\mathbf{t}), \dots, w_N(\mathbf{t})]'$$

where

$$w_i(\mathbf{t}) = (y_i - g(x_i, \hat{\mathbf{a}}(\mathbf{t}), \mathbf{t})) / \sigma_i.$$

We solve the nonlinear least-squares problem

$$\min_{\mathbf{t}} W(\mathbf{t}) = \min_{\mathbf{t}} \frac{1}{2} \mathbf{w}(\mathbf{t}) \mathbf{w}'(\mathbf{t}) \quad (\text{A3})$$

by a modified version of NOLLS1.<sup>17</sup> Let the Jacobian matrix of  $\mathbf{w}(\mathbf{t})$  be denoted by

$$J(\mathbf{t}) = \begin{bmatrix} \partial w_1 / \partial t_1 & \cdot & \cdot & \partial w_1 / \partial t_M \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \partial w_N / \partial t_1 & \cdot & \cdot & \partial w_N / \partial t_M \end{bmatrix}.$$

Then, the gradient  $\nabla W$  of  $W$  is given by

$$\nabla W(\mathbf{t}) = J'(\mathbf{t}) \mathbf{w}(\mathbf{t})$$

and the Hessian matrix  $\nabla^2 W$  is given by

$$\nabla^2 W(\mathbf{t}) = J'(\mathbf{t}) J(\mathbf{t}) + \sum_{i=1}^N w_i(\mathbf{t}) \nabla^2 w_i(\mathbf{t}). \quad (\text{A4})$$

The Newton-Raphson method<sup>28</sup> for solving the problem is written in terms of the following iterative process: Starting from an initial guess  $\mathbf{t}^{(0)}$ , we generate the sequence of  $\mathbf{t}$  by the formula,

$$\mathbf{t}^{(k+1)} = \mathbf{t}^{(k)} - [\nabla^2 W(\mathbf{t}^{(k)})]^{-1} \nabla W(\mathbf{t}^{(k)}). \quad (\text{A5})$$

Analytic formulas for  $J(\mathbf{t})$ ,  $\nabla W(\mathbf{t})$ , and  $\nabla^2 W(\mathbf{t})$  are not available in our case. A finite difference approximation of  $J(\mathbf{t})$  is actually used to calculate  $\nabla W(\mathbf{t})$  and  $J'(\mathbf{t}) J(\mathbf{t})$ . In quasi-Newton methods, particular approximations of  $\nabla^2 W(\mathbf{t})$  are used in place of  $\nabla^2 W(\mathbf{t})$  itself, which are iteratively corrected by update formulas. However, since the first term in the right-hand side of (A4) is already available, we only have to estimate the second term,

$$S(\mathbf{t}) = \sum_{i=1}^N w_i(\mathbf{t}) \nabla^2 w_i(\mathbf{t}). \quad (\text{A6})$$

We used the symmetric single rank update formula

$$S^{(k+1)} = S^{(k)} + \frac{(\mathbf{y} - S^{(k)} \mathbf{s})(\mathbf{y} - S^{(k)} \mathbf{s})'}{(\mathbf{y} - S^{(k)} \mathbf{s})' \mathbf{s}} \quad (\text{A7})$$

or the Broyden,<sup>29</sup> Fletcher,<sup>30</sup> Goldfarb,<sup>31</sup> and Shanno<sup>32</sup> formula

$$S^{(k+1)} = S^{(k)} + \frac{\mathbf{y} \mathbf{y}'}{\mathbf{y}' \mathbf{s}} - \frac{\mathbf{H} \mathbf{s} \mathbf{s}' \mathbf{H}}{\mathbf{s}' \mathbf{H} \mathbf{s}} \quad (\text{A8})$$

to approximate  $S(\mathbf{t})$ , where Bigg's formulas

$$\mathbf{s} = \mathbf{t}^{(k)} - \mathbf{t}^{(k-1)} \quad (\text{A9})$$

and

$$\mathbf{y} = [J(\mathbf{t}^{(k)}) - J(\mathbf{t}^{(k-1)})]' \mathbf{w}(\mathbf{t}^{(k)}) \quad (\text{A10})$$

are used in our version of NOLLS1.

## APPENDIX B: MINIMUM AKAIKE'S INFORMATION CRITERION PROCEDURE

The minimum AIC procedure is a statistical method which was introduced by Akaike<sup>21,22</sup> to select the best of the alternative parameter models based on the statistical data. The procedure is interpreted as a maximization of the expected entropy of the predictive distribution. The entropy, or negative sign of Kullback-Leibler information,  $-I(p; q)$ , or the unknown true distribution with respect to the fitted distribution is given by

$$\begin{aligned} B(p; q) &= - \int \frac{p(z)}{q(z|\mathbf{a})} \ln \frac{p(z)}{q(z|\mathbf{a})} q(z|\mathbf{a}) dz \\ &= \int p(z) \ln \frac{q(z|\mathbf{a})}{p(z)} dz \\ &= E_z \ln \frac{q(z|\mathbf{a})}{p(z)} \\ &= E_z \ln q(z|\mathbf{a}) - E_z \ln p(z), \end{aligned} \quad (\text{B1})$$

where  $p(z)$  and  $q(z|\mathbf{a})$  are density functions of true and fitted distributions respectively,  $\mathbf{a}$  denotes the adjustable parameters in the fitted model and  $E_z$  denotes the mean or expectation with respect to  $z$ . Since

$$B(p; q) \leq 0$$

and the equality holds if and only if

$$q(z|\mathbf{a}) = p(z), \text{ almost everywhere.}$$

$B(p; q)$  can be considered as the measure of the deviation of  $q(z|\mathbf{a})$  from  $p(z)$ . As  $E_z \ln p(z)$  is independent of  $q(z|\mathbf{a})$ , it follows that the larger value of  $E_z \ln q(z|\mathbf{a})$  denotes  $q(z|\mathbf{a})$  closer to  $p(z)$ . Goodness of fitting  $q(z|\mathbf{a})$  to the unknown  $p(z)$  can thus be evaluated in terms of the value  $E_z \ln q(z|\mathbf{a})$  based on the data.

Let  $z_1, z_2, \dots, z_N$  be the observed data which are governed by the true distribution  $p(z)$ , and assume  $q(z|\mathbf{a})$  a parametric density model with parameter  $\mathbf{a}$ . Then the log-likelihood  $l(\mathbf{a})$  of the parameter  $\mathbf{a}$  is given by

$$l(\mathbf{a}) = \sum_{i=1}^N \ln q(z_i|\mathbf{a}) \quad (\text{B2})$$

and  $l(\mathbf{a})/N$  forms a natural estimate of  $E_z \ln q(z|\mathbf{a})$  where  $\mathbf{a} = \hat{\mathbf{a}}$  is estimated from the data. If we use  $l(\hat{\mathbf{a}})/N$  as an estimate of  $E_z \ln q(z|\mathbf{a})$ , it is biased asymptotically by the amount  $-2k$ , where  $k$  is the number of free parameters in  $\mathbf{a}$ .

By compensating this bias, Akaike obtained the AIC statistics,

$$AIC(\hat{a}|z_1, z_2, \dots, z_N) = 2l(\hat{a}) + 2\kappa,$$

where  $\hat{a}$  is the maximum likelihood estimate of  $a$ . Among the alternative models  $[q(z|a)]$ , the minimum AIC procedure chooses the one for which the value of the AIC statistics is the smallest.

- <sup>1</sup>H. C. Burger and P. H. Van Cittert, *Z. Phys.* **79**, 722 (1932).
- <sup>2</sup>O. Kratky, G. Porod, and Z. Skala, *Acta Phys. Austriaca* **13**, 76 (1960).
- <sup>3</sup>J. Schelten and F. Hossfeld, *J. Appl. Cryst.* **4**, 210 (1971).
- <sup>4</sup>I. Isenberg, *J. Chem. Phys.* **59**, 5696 (1973).
- <sup>5</sup>I. Isenberg, *J. Chem. Phys.* **59**, 5708 (1973).
- <sup>6</sup>D. V. O'Conner, W. R. Ware, and J. C. Andre, *J. Phys. Chem.* **83**, 1333 (1979).
- <sup>7</sup>E. Freire and R. L. Biltonen, *Biopolymers* **17**, 463 (1978).
- <sup>8</sup>E. Freire and R. L. Biltonen, *Biopolymers* **17**, 481 (1978).
- <sup>9</sup>E. Freire and R. L. Biltonen, *Biopolymers* **17**, 497 (1978).
- <sup>10</sup>J. A. Lake, *Acta Crystallogr.* **19**, 938 (1967).
- <sup>11</sup>See, for example, J. C. Andre, L. M. Vincent, D. V. O'Conner and W. R. Ware, *J. Phys. Chem.* **83**, 2285 (1979).
- <sup>12</sup>O. Glatter, *J. Appl. Cryst.* **10**, 415 (1977).
- <sup>13</sup>K. Ichida and F. Yoshimoto, in *Spline Kansu to Sono Oyo: Spline Functions and Their Applications* (Kyoiku Shuppan, Tokyo, 1979).
- <sup>14</sup>L. L. Schumaker, in *Theory and Applications of Spline Functions*, edited by T. N. E. Greville (Academic, New York, 1969), pp. 65–102.
- <sup>15</sup>K. Tanabe, *Oyo Tokeigaku, Appl. Stat.* **9**, 119 (1981).
- <sup>16</sup>M. C. Barthoromew-Biggs, *Math. Programming* **12**, 67 (1977).
- <sup>17</sup>K. Tanabe and S. Ueda, in *Computer Science Monograph No. 17*, edited by M. Ishida (Institute of Statistical Mathematics, Tokyo, 1981).
- <sup>18</sup>H. Akaike, in *Proc. 3rd Hawaii Intl. Conf. on System Sciences* (Univ. of Hawaii, Honolulu, 1970), pp. 974–977.
- <sup>19</sup>H. Akaike, *IEEE Trans. Automation Control*, **AC-19**, 716 (1974).
- <sup>20</sup>H. Akaike, *Suri Kagaku Math. Sci.* **153**, 5 (1976).
- <sup>21</sup>H. Akaike, in *Proc. 2nd Intl. Symp. on Information Theory*, edited by B. N. Petrov and F. Csaki (Akademiai Kiado, Budapest, 1973).
- <sup>22</sup>H. Akaike, in *Application of Statistics*, edited by P. K. Krishniah (North Holland, Amsterdam, 1977), pp. 27–41.
- <sup>23</sup>K. Tanabe, *Suri Kagaku Math. Sci.* **153**, 1 (1976).
- <sup>24</sup>G. Wahba and S. Wold, *Commun. Stat.* **4**, 1 (1975).
- <sup>25</sup>O. Kratky and Z. Skala, *Z. Elektrochem.* **62**, 73 (1958).
- <sup>26</sup>V. Luzzati and R. Baro, *J. Phys. Radiat. Phys. Appl. Suppl.* **17**, 186 A (1961).
- <sup>27</sup>O. Glatter, *Acta Phys. Austriaca* **47**, 83 (1977).
- <sup>28</sup>See, for example, J. Kowalik and M. R. Osborne, in *Method for Unconstrained Optimization Problems* (Elsevier, New York, 1968).
- <sup>29</sup>C. G. Broyden, *Math. Comp.* **21**, 368 (1969).
- <sup>30</sup>R. Fletcher, *Computer J.* **13**, 317 (1970).
- <sup>31</sup>G. H. Goldfarb, in *Optimization*, edited by R. Fletcher (Academic, London, 1969), pp. 273–281.
- <sup>32</sup>D. F. Shanno, *Math. Comp.* **24**, 647 (1970).

# Whole mantle P-wave travel time tomography

Hiroshi Inoue

*International Institute of Seismology and Earthquake Engineering, Building Research Institute, Tsukuba, Ibaraki 305 (Japan)*

Yoshio Fukao

*Department of Earth Sciences, Nagoya University, Nagoya (Japan)*

Kunio Tanabe and Yoshihiko Ogata

*The Institute of Statistical Mathematics, Minami Azabu, Tokyo (Japan)*

(Received September 16, 1988; revised July 17, 1989; accepted August 9, 1989)

Inoue, H., Fukao, Y., Tanabe, K. and Ogata, Y., 1990. Whole mantle P-wave travel time tomography. *Phys. Earth Planet. Inter.*, 59: 294–328.

A method of tomographic inversion to obtain three-dimensional velocity perturbations in the Earth's whole mantle has been developed, and applied to more than two million P-wave arrival time data reported by International Seismology Center (ISC). The model is parameterized with 32,768 blocks; the divisions in latitude, longitude, and radius are 32, 64, and 16, respectively. Horizontal cell size is  $5.6^\circ \times 5.6^\circ$ . The layer thicknesses vary with depth: 29 km just below the surface and 334 km just above the core–mantle boundary. Starting from a spherically symmetric Earth model, we obtained a three-dimensional model using the following iterative procedures. First, we relocated all the events; second, we backprojected the residual into the whole mantle; third, we refined the spherically symmetric Earth model. The solutions have been converged in five iterations. We adopted the following techniques in the backprojection procedure. The first order smoothness was introduced as a damping, which makes the solution independent of the starting model and its apparent fluctuation minimal. The basic equations for delay times and smoothness are solved using the conjugate gradient method, an iterative method which guarantees the convergence of solution into the exact least squares solution. The weight on the smoothness, i.e., the damping factor, was objectively determined by a simplified cross validation technique. The final solution was obtained as an average of the ten solutions, each of which was derived from one tenth of the total data set. The reliability of the solution is examined in two ways: (1) mapping the resolution given by the reconstruction of checkerboard patterns, and (2) mapping the variance given from the gaussian noise input.

This is the first result which delineates the whole mantle structure with its reliability mapping. The longest wavelength anomalies of the lower mantle are similar to those of previous studies by Dziewonski (1984) and Hager and Clayton (1989). However, the shorter wavelength patterns have many discrepancies. The fast anomalies of the uppermost mantle beneath the shield and the slow anomalies beneath the active tectonic regions are clearly seen. We also revealed the  $l = 2$  pattern at the transition zone reported by the previous surface wave and free oscillation studies. The most striking feature is that of the fast anomalies lying along the extension of the large subducting plates up to depths of about 1000 km or more. The resolving powers in these regions are generally good.

## 1. Introduction

Imaging the lateral inhomogeneity of the Earth's deep interior is the most fascinating subject of

present day seismology, and is of primary importance for understanding the global tectonics and the dynamic evolution of our planet Earth. Since the late 1960's, many seismologists have made

continuous efforts in attempting to map the aspherical structure more clearly and accurately. In the 1980's, increasing quality and quantity of global data set and large computer powers have opened a field of seismology called "Seismic Tomography". Tomography inversion of surface waves has revealed the several percent of S-wave velocity perturbations in the upper mantle on a global scale (Nakanishi and Anderson, 1984; Woodhouse and Dziewonski, 1984; Nataf et al., 1986). There have been also regional scale surface wave studies that are intended to increase a spatial resolution at the expense of restricting the region of analysis to a part of the sphere (Montagner, 1986; Suetsugu and Nakanishi, 1987; Honda and Tanimoto, 1987). There has also been an effort to map a deeper part of the mantle (Tanimoto, 1988). Further deep interiors of the Earth have been mapped using free oscillation spectra (Giardini et al., 1987).

Body wave travel time tomography, on the other hand, gives information independent of the surface wave and free oscillation studies. Since a large number of rays pass through deep mantle, they can provide images of high frequency P- and S-velocity perturbations with high spatial resolution. The pioneer works of the travel time tomography are Aki et al. (1977) for a regional scale, and Dziewonski et al. (1977) for the global scale. The latter obtained the lower mantle P-wave velocity heterogeneity using the vast amount of travel time data compiled by International Seismology Center (ISC). Their work was followed by the worldwide station anomaly study of Dziewonski and Anderson (1983) and then by Dziewonski (1984) who constructed a lower mantle P-velocity model of harmonic expansion up to degree and order 6. They, and also Morelli and Dziewonski (1987), adopted the approach of expressing the Earth structure by a harmonic expansion and obtained lower mantle velocity perturbations of up to  $\pm 1.5\%$ . Clayton and Comer (1983) and Hager and Clayton (1989), on the other hand, presented the 3-D model of the lower mantle expanded by a number of blocks, using the similar large data set compiled by ISC. Their result shows  $\pm 1\%$  perturbations after smoothing. Both Dziewonski (1984) and Hager and Clayton (1989) re-

constructed the lower mantle structure by reducing the upper mantle anomalies to the source and station anomalies.

Although there are a number of local tomographic studies of the upper few hundred kilometers of the mantle from P time data, we have not yet been given an image of the 3-D heterogeneity of the whole mantle. The tomographic inversion presented in this paper is the first attempt to obtain such an image with high resolution and high accuracy, using both shallow and deep events and both near and distant P phases, together with mapping its reliability. The main purpose of this paper is to describe our method of inversion and show how well it works in application to real data set.

## 2. Method

First we summarize our method of tomographic inversion. The data set is of worldwide P-wave travel times for a number of shallow to deep events reported by ISC. The events are selected carefully so as to be distributed as uniformly as possible. The detailed selection scheme is described later. The three-dimensional slowness perturbation is modeled by a number of blocks. The inversion is performed in the following iterative algorithm.

- (1) Start from an initial model;
- (2) repeat until convergence: (a) relocation of events; (b) backprojection to the 3-D slowness perturbation; (c) refinement of the spherically symmetric model;
- (3) calculation of the resolution and variance.

The unknown parameters to be determined are the spherically symmetric velocity model, 3-D slowness perturbations, and hypocentral parameters of all the events used. No station anomalies are dealt with; if they exist they are addressed to the perturbations in the uppermost structure. In the following subsections, we describe each part of the inverting processes in detail.

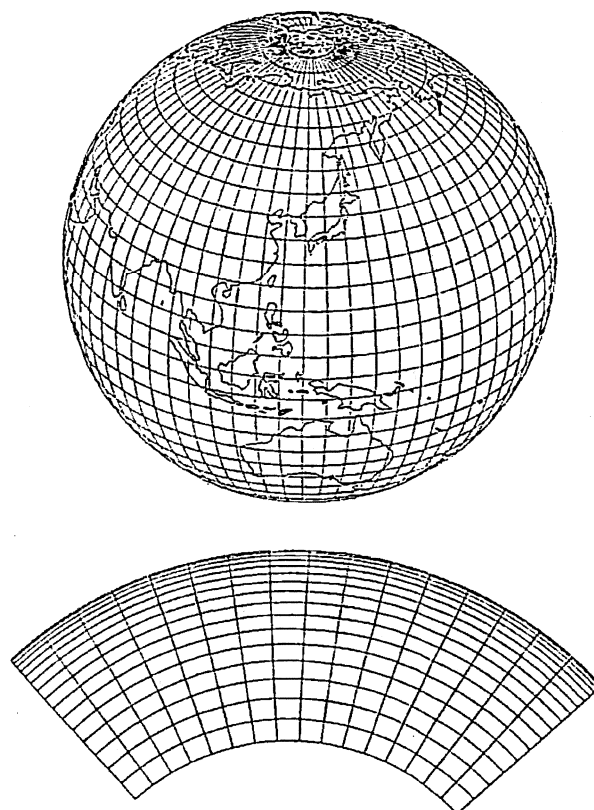
### 2.1. Model

The P-wave velocity distribution in the mantle is expressed as a sum of the spherically averaged

velocity–depth function and the 3-D perturbation on it. Dziewonski (1984) and Morelli and Dziewonski (1987) represented the 3-D structure of the Earth with the spherical harmonic expansion, rather than with blocks as used by Clayton and Comer (1983) and Hager and Clayton (1988). Their reasoning of the use of spherical harmonics are as follows: (1) the block parameterization requires more parameters to express the model of the same level of smoothness realized with harmonic expansion; (2) it is impractical to estimate the resolution and variance of the model for such large numbers of parameters; and (3) the harmonic model can be compared directly with other geophysical maps such as gravity and magnetic fields. We, however, point out the following. (1) Capability of representing a distribution by blocks is essentially the same as representing one by harmonics. If one would like to avoid sharp edges of the block model in the final presentation, he may smooth out the edges using an appropriate filter, or may adopt an expansion with higher order spline basis. (2) Estimates of resolutions and variances are not impractical in the block model with such a large number of parameters as used in this paper. (3) Construction of a spherical harmonic model from the block model is easy.

In our view, the distribution in the Earth's interior must be described with a system of local basis, such as blocks or higher order splines. This is because, firstly, the solution is always biased due to an uneven distribution of the data. Of course the truncated expansion of the Earth's lateral heterogeneity by either harmonics or blocks is only an approximation. The bias associated with the local basis is, however, still smaller than the harmonics. A local basis approach is more appropriate because, secondly, it has a greater numerical advantage over a global basis approach for a problem with a large number of model parameters. In the latter approach, the nature of full coefficient matrix of the basic observational equation makes the inversion formidable. For the above two reasons, we adopted the model expanded by block basis.

The configuration of the blocks in our inversion is shown in Fig. 1, where the divisions in latitude, longitude and depth are 32, 64, and 16,



**$32 \times 64 \times 16 = 32768$  blocks**

Fig. 1. Block configuration of the model. Upper, surface view; lower, vertical cross section along the equator.

respectively. The total number of blocks is 32,768. The surface cell size is  $5.625^\circ \times 5.625^\circ$ . The depth division varies with depth: the thickness of the shell is smaller in the shallower depth (Fig. 1 lower). The depth nodes  $h_i$  are determined from the quadratic formula

$$h_i = \frac{2600}{16^2} i^2 + \frac{300}{16} i \quad i = 0 \dots 16 \quad (1)$$

from the surface ( $i = 0$ ) to the core mantle boundary (CMB) ( $i = 16$ ) at 2900 km. The thickness of the shells are 28.9 km just below the surface, 171 km just above the depth of 800 km, and 334 km just above the CMB. Such non-uniform division is made to model the greater complexity at shallower depths, because we know a priori that the velocity variation in either the spherically averaged model or its lateral perturbation is greater in the shallower depths. The ray

paths are computed in the spherically symmetric model and a path segment in a block is assumed to be a straight line.

## 2.2. Relocation

All the events are relocated one by one using a standard algorithm. First we construct the travel time table  $T(\Delta, h)$  for the spherically averaged  $V_p$  model given in the preceding iteration, where  $\Delta$  is the epicentral distance and  $h$  is the hypocentral depth. Linearized relocation equations for an observed arrival time  $t_{\text{obs}}$  is

$$\frac{\partial t}{\partial \theta} \delta \theta + \frac{\partial t}{\partial \phi} \delta \phi + \frac{\partial t}{\partial h} \delta h + \delta t_0 = t_{\text{obs}} - (t_0 + t_{1D} + t_{3D}) \quad (2)$$

where  $\theta$ ,  $\phi$ ,  $h$ ,  $t_0$  are latitude, longitude, depth, and origin time of the event, respectively.  $t_{1D}$  is the travel time calculated from the table  $T(\Delta, h)$  of the spherically averaged model.  $t_{3D}$  is a perturbation due to 3-D slowness anomaly given by

$$t_{3D} = \sum_j l_j s_j \quad (3)$$

where  $l_j$  is the length of the ray segment in the  $j$ -th block and  $s_j$  is the slowness perturbation of the  $j$ -th block determined in the preceding backprojection in the iterative process. A set of eqns. (2) for an event are solved by the least squares method. We assume that the standard deviations of the observational errors are 0.1, 0.3, 1 s for readings  $iP$ ,  $P$ , and  $eP$  in the ISC's notation respectively. Since the data sometimes contain large errors not following the gaussian distribution, we adopted a robust estimate by applying weight  $w$  as a function of residual  $r$  as

$$w(r) = \begin{cases} [1 - (r/10)^2]^2, & |r| < 10s \\ 0, & |r| \geq 10s \end{cases} \quad (4)$$

This weighting scheme is also applied in the backprojection process which will be explained later. Since the relocation is a non-linear process, we iterated the least squares calculation ten times introducing a damping. The calculation of  $t_{3D}$  is made only at the first iteration because it is very time consuming.

Some events, such as those along mid-oceanic ridges, are poorly located because of a trade-off between focal depth and origin time. The depth of such an event is limited to a reasonable range according to the tectonic province. The range is 3–20 km at mid-oceanic ridges, 0–700 km at subduction zones, 0–100 km at continents and 3–50 km at oceans. A structure in a region having such unstable events may not be well resolved. The effect of mislocation, however, does not significantly affect the structure of the distant regions from a source.

## 2.3. Backprojection

The travel time residuals after relocation are backprojected into the mantle. The equation for the backprojection of a residual  $t_i$  is

$$t_i = \sum_j l_{ij} s_j \quad (5)$$

The set of eqn. (5) is written in a matrix form as  $\mathbf{d} = \mathbf{G}\mathbf{m}$  (6)

where  $\mathbf{d}$  is the residual vector,  $\mathbf{G}$  is the data kernel, and  $\mathbf{m}$  is the unknown model (slowness perturbation) vector. The system of eqn. (6) is normally both inconsistent and underdeterminant, so we solve it by the least squares method introducing certain a priori information into the solution rather than the data.

There are many types of inversion of eqn. (6). The most popular one is the damped least squares, which was first applied in travel time inversion by Aki et al. (1977) and followed by many authors. The solution of the damped least squares is given by solving the normal equation

$$(\mathbf{G}^T \mathbf{G} + \theta^2 \mathbf{I}) \mathbf{m} = \mathbf{G}^T \mathbf{d} \quad (7)$$

where the term  $\theta^2 \mathbf{I}$  is introduced to stabilize the solution (Franklin, 1970; Jackson, 1979).  $\theta$  is a damping constant determined a priori. The a priori knowledge is that the model  $\mathbf{m}$  is a gaussian noise of zero mean, which requires the solution to be as close to the starting model as possible if it is not properly constrained by the data. Equation (7) is usually solved using the Cholesky's decomposition, which is a direct elimination method suitable for a system of equations with a symmetric matrix.



In this study, we do not use this method because of the following two reasons. First, both the construction of the normal equation and solving it using a direct elimination method are impractical since the number of model parameters is extremely large in our problem. Recent studies of large scale tomographic inversion do not adopt the damped least-squares in this sense (Humphreys et al., 1984; Nakanishi, 1985; Grand, 1987; Suetsugu and Nakanishi, 1987; Hager and Clayton, 1989; and Humphreys and Clayton, 1988). Second, the a priori knowledge of  $m$  = starting model + gaussian error does not seem to be a correct one. Alternatively, first, we adopt the conjugate gradient method to solve the normal equation and, second, introduce the first order smoothness prior.

### 2.3.1. Smoothness prior

An introduction of smoothness to a problem of modeling with insufficient data has been made frequently. The simplest example is the line graph. The cubic spline fitting is also widely known, which requires the minimal curvature of the model in the least-squares sense. First we explain the smoothness prior to the 1-dimensional point data fitting.

Suppose that we are required to fit a model function  $y = f(x)$  to observed data points  $(x_i, y_i)$   $i = 1 \dots N$ . The solution which we would like to obtain is such that misfits least to the data and that is smooth in the least squares sense. The solution can be given by a minimization of the following  $l_2$  norm

$$\Pi = \sum_{i=1}^N [y_i - f(x_i)]^2 + \lambda \int_{\Omega} [f^{(k)}(x)]^2 dx \quad (8)$$

The first and the second terms of the right hand side (RHS) represent the misfit to the data and the roughness of the model, respectively, where  $f^{(k)}$  denotes the  $k$ -th order derivative of  $f$ ,  $\Omega$  is the domain of  $x$  under consideration, and  $\lambda$  is an arbitrary weighting factor on the roughness norm. The second term may be a summation of roughness norms with different orders. The estimate  $f$  which minimizes  $\Pi$  is the best fit and the smoothest function in the admissible functional

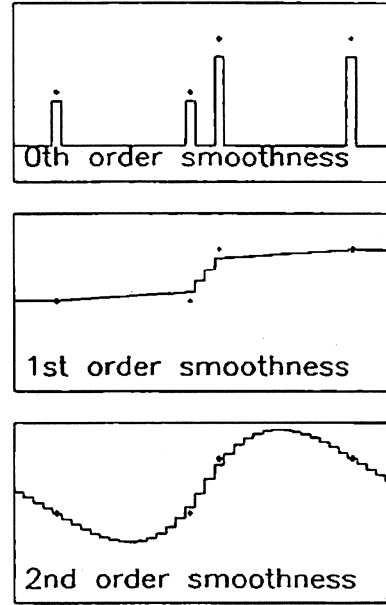


Fig. 2. Role of three different orders of smoothness introduced in 1-D point fitting problem. Upper, ordinary damped least squares; middle, order adopted in this paper.

space. The a priori information used here is that the  $k$ -th order derivative of the function  $f$  is zero everywhere plus gaussian noise. The criterion with  $k = 0$  is the standard damped least-squares.

Figure 2 illustrates the role of the zero, the first and the second order smoothnesses in fitting problem. The solution with the zero order smoothness depends on the initial solution while those with the first and higher order smoothnesses do not. A physical model to fit the first or the second order smoothness is a string under tension or an elastic bar free from tension, respectively. Detail of the fitting problem in one or two dimensions is described in Inoue (1986).

We adopted the first order smoothness to our problem of tomography, because of the minimal oscillatory nature of the solution and its independence of the initial solution. The former is an important property for suppressing artificial fluctuation of the solution, since we sometimes pay attention to the change of the physical property of the Earth's interior as well as to the value itself. The  $l_2$  norm to be minimized in our global

tomography for a continual slowness perturbation  $s(r, \theta, \phi)$  is written as

$$\Pi = \sum_i \frac{1}{\sigma_i^2} \left( t_i - \int_{\Gamma_i} s d\Gamma \right)^2 + \frac{1}{l_u^3} \int_{V_M} \left\{ \frac{1}{\sigma_h^2 r^2} \left[ \left( \frac{\partial s}{\partial \theta} \right)^2 + \frac{1}{\sin^2 \theta} \left( \frac{\partial s}{\partial \phi} \right)^2 \right] + \frac{1}{\sigma_v^2} \left( \frac{\partial s}{\partial r} \right)^2 \right\} dV \quad (9)$$

The first and the second terms express the misfit and the first order roughness respectively.  $\sigma_i$  is the standard error of the observation  $t_i$  and  $\sigma_h$  and  $\sigma_v$  are the a priori standard errors of the horizontal and the vertical smoothnesses respectively, which will be determined by the cross validation described later. In this paper, we refer to  $\sigma_h$  and  $\sigma_v$  as the roughness parameters. The integral in the first term is along the ray  $\Gamma_i$  and that in the second term is over entire volume of the mantle  $V_M$ . The damping factors corresponding to  $\lambda$  in eqn. (8) are  $l_u^{-3} \sigma_h^{-2}$  and  $l_u^{-3} \sigma_v^{-2}$ , where  $l_u$  is a unit length introduced to make the derivative norm dimensionless, which is only for convenience and has no essential meaning. The interpretation of  $l_u$  is given in Inoue (1986) in detail. The horizontal and vertical smoothnesses are dealt with separately, since the Earth is intrinsically vertically heterogeneous.

Rewriting eqn. (9) to a discrete model, we have

$$\Pi = \sum_i \frac{1}{\sigma_i^2} \left( t_i - \sum_j l_{ij} s_j \right)^2 + \frac{1}{l_u^3} \sum_V \left[ \frac{1}{\sigma_h^2} \left( \frac{\Delta S_\theta}{r \Delta \theta} \Delta s_\theta^2 + \frac{\Delta S_\phi}{r \Delta \phi \sin \theta} \Delta s_\phi^2 \right) + \frac{1}{\sigma_v^2} \frac{\Delta S_r}{\Delta r} \Delta s_r^2 \right] \quad (10)$$

where  $\Delta \theta$ ,  $\Delta \phi$  and  $\Delta r$  are the step widths and  $\Delta S_\theta$ ,  $\Delta S_\phi$  and  $\Delta S_r$  are the contact areas of neighbouring blocks along the coordinates, which are given by  $r \Delta \phi \Delta r \sin \theta$ ,  $r \Delta \theta \Delta r$  and  $r^2 \Delta \theta \Delta \phi \sin \theta$ , respectively.  $\Delta s_\theta$ ,  $\Delta s_\phi$  and  $\Delta s_r$  are the differences of the neighbouring model parameters along the coordinates.

The system of equation to be solved in the least squares method is

$$\begin{pmatrix} G \\ D \end{pmatrix} m = \begin{pmatrix} d \\ 0 \end{pmatrix} \quad (11a)$$

with the covariance matrix in the form

$$\begin{pmatrix} C_d & 0 \\ 0 & C_m \end{pmatrix} \quad (11b)$$

where  $C_d$  is the covariance matrix of the data and  $C_m$  is the covariance matrix of the smoothness, of which diagonal elements are  $\sigma_h^2$  and  $\sigma_v^2$  and all of the off-diagonals are zeros. The smoothness kernel  $D$  is composed of the coefficients for the first derivatives, whose row vector is  $(\dots, c, -c, \dots)$ . The coefficients  $c$  for the three directions of the coordinates  $\theta$ ,  $\phi$ ,  $r$  are  $\sqrt{\Delta S_\theta / l_u^3 r \Delta \theta}$ ,  $\{\Delta S_\phi / l_u^3 r \Delta \phi \sin \theta\}^{1/2}$  and  $\sqrt{\Delta S_r / l_u^3 \Delta r}$ , respectively.

In each step in the iterative process of the relocation and the backprojection, eqn. (11a) must be modified to

$$\begin{pmatrix} G \\ D \end{pmatrix} \delta m = \begin{pmatrix} \delta d \\ -D m_0 \end{pmatrix} \quad (12)$$

where  $m_0$  is the starting model. The normal equation for eqn. (12) is given by

$$\begin{aligned} (G^T C_d^{-1} G + D^T C_m^{-1} D) \delta m \\ = G^T C_d^{-1} \delta d - D^T C_m^{-1} D m_0 \end{aligned} \quad (13)$$

The solution of the normal equation can be obtained by an iterative method without constructing eqn. (13) explicitly.

The a priori covariance matrix  $C_m$  of gaussian type, which was proposed by Tarantola and Valette (1982), is sometimes used for providing a smoothness of the solution. However, its basic property is the same as the standard damped least squares plus a posteriori moving average filtering. In their framework, our smoothness is formally given by the a priori covariance  $(D^T D)^{-1}$ , which is the covariance function of the inverse of the second order differential operator  $(\dots, -1, 2, -1, \dots)$ , that is, the second order integration operator.

### 2.3.2. Conjugate gradient method

Recent works on large scale travel time tomography adopt iterative algorithms to solve the equation system, because the iterative scheme is very suitable for solving a system of equations with large and sparse coefficient matrix. There have been many theoretical and experimental works on the comparison of backprojection techniques

tomography for a continual slowness perturbation  $s(r, \theta, \phi)$  is written as

$$\Pi = \sum_i \frac{1}{\sigma_i^2} \left( t_i - \int_{\Gamma_i} s d\Gamma \right)^2 + \frac{1}{l_u^3} \int_{V_M} \left\{ \frac{1}{\sigma_h^2 r^2} \left[ \left( \frac{\partial s}{\partial \theta} \right)^2 + \frac{1}{\sin^2 \theta} \left( \frac{\partial s}{\partial \phi} \right)^2 \right] + \frac{1}{\sigma_v^2} \left( \frac{\partial s}{\partial r} \right)^2 \right\} dV \quad (9)$$

The first and the second terms express the misfit and the first order roughness respectively.  $\sigma_i$  is the standard error of the observation  $t_i$  and  $\sigma_h$  and  $\sigma_v$  are the a priori standard errors of the horizontal and the vertical smoothnesses respectively, which will be determined by the cross validation described later. In this paper, we refer to  $\sigma_h$  and  $\sigma_v$  as the roughness parameters. The integral in the first term is along the ray  $\Gamma_i$  and that in the second term is over entire volume of the mantle  $V_M$ . The damping factors corresponding to  $\lambda$  in eqn. (8) are  $l_u^{-3} \sigma_h^{-2}$  and  $l_u^{-3} \sigma_v^{-2}$ , where  $l_u$  is a unit length introduced to make the derivative norm dimensionless, which is only for convenience and has no essential meaning. The interpretation of  $l_u$  is given in Inoue (1986) in detail. The horizontal and vertical smoothnesses are dealt with separately, since the Earth is intrinsically vertically heterogeneous.

Rewriting eqn. (9) to a discrete model, we have

$$\Pi = \sum_i \frac{1}{\sigma_i^2} \left( t_i - \sum_j l_{ij} s_j \right)^2 + \frac{1}{l_u^3} \sum_V \left[ \frac{1}{\sigma_h^2} \left( \frac{\Delta S_\theta}{r \Delta \theta} \Delta s_\theta^2 + \frac{\Delta S_\phi}{r \Delta \phi \sin \theta} \Delta s_\phi^2 \right) + \frac{1}{\sigma_v^2} \frac{\Delta S_r}{\Delta r} \Delta s_r^2 \right] \quad (10)$$

where  $\Delta \theta$ ,  $\Delta \phi$  and  $\Delta r$  are the step widths and  $\Delta S_\theta$ ,  $\Delta S_\phi$  and  $\Delta S_r$  are the contact areas of neighbouring blocks along the coordinates, which are given by  $r \Delta \phi \Delta r \sin \theta$ ,  $r \Delta \theta \Delta r$  and  $r^2 \Delta \theta \Delta \phi \sin \theta$ , respectively.  $\Delta s_\theta$ ,  $\Delta s_\phi$  and  $\Delta s_r$  are the differences of the neighbouring model parameters along the coordinates.

The system of equation to be solved in the least squares method is

$$\begin{pmatrix} G \\ D \end{pmatrix} m = \begin{pmatrix} d \\ 0 \end{pmatrix} \quad (11a)$$

with the covariance matrix in the form

$$\begin{pmatrix} C_d & 0 \\ 0 & C_m \end{pmatrix} \quad (11b)$$

where  $C_d$  is the covariance matrix of the data and  $C_m$  is the covariance matrix of the smoothness, of which diagonal elements are  $\sigma_h^2$  and  $\sigma_v^2$  and all of the off-diagonals are zeros. The smoothness kernel  $D$  is composed of the coefficients for the first derivatives, whose row vector is  $(\dots, c, -c, \dots)$ . The coefficients  $c$  for the three directions of the coordinates  $\theta$ ,  $\phi$ ,  $r$  are  $\sqrt{\Delta S_\theta / l_u^3 r \Delta \theta}$ ,  $\{\Delta S_\phi / l_u^3 r \Delta \phi \sin \theta\}^{1/2}$  and  $\sqrt{\Delta S_r / l_u^3 \Delta r}$ , respectively.

In each step in the iterative process of the relocation and the backprojection, eqn. (11a) must be modified to

$$\begin{pmatrix} G \\ D \end{pmatrix} \delta m = \begin{pmatrix} \delta d \\ -D m_0 \end{pmatrix} \quad (12)$$

where  $m_0$  is the starting model. The normal equation for eqn. (12) is given by

$$(G^T C_d^{-1} G + D^T C_m^{-1} D) \delta m = G^T C_d^{-1} \delta d - D^T C_m^{-1} D m_0 \quad (13)$$

The solution of the normal equation can be obtained by an iterative method without constructing eqn. (13) explicitly.

The a priori covariance matrix  $C_m$  of gaussian type, which was proposed by Tarantola and Valette (1982), is sometimes used for providing a smoothness of the solution. However, its basic property is the same as the standard damped least squares plus a posteriori moving average filtering. In their framework, our smoothness is formally given by the a priori covariance  $(D^T D)^{-1}$ , which is the covariance function of the inverse of the second order differential operator  $(\dots, -1, 2, -1, \dots)$ , that is, the second order integration operator.

### 2.3.2. Conjugate gradient method

Recent works on large scale travel time tomography adopt iterative algorithms to solve the equation system, because the iterative scheme is very suitable for solving a system of equations with large and sparse coefficient matrix. There have been many theoretical and experimental works on the comparison of backprojection techniques

given by the incomplete Cholesky decomposition. In the conjugate gradient algorithm shown above, it is not required to calculate  $L^{-T}$  explicitly. Since system (18) is better conditioned than system (14), the solution converges faster. However, the incomplete Cholesky decomposition requires rather complicated algorithm, so that we adopted the simplest version of the preconditioning, i.e., the diagonal scaling. In the diagonal scaling algorithm,  $L^{-T}$  is only a diagonal matrix given by

$$L^{-T} = [\text{diag}(A^T A)]^{-1/2} \quad (20)$$

which is applied in eqns. (18) and (19) explicitly. The scaling does not affect the relative weight on each equation.

### 2.3.3. Cross validation

The standard errors of the smoothness  $\sigma_h$  and  $\sigma_v$  (the roughness parameters) must be determined a priori. Unfortunately, we know nothing about them a priori. One method to determine the parameters is ABIC (Akaike's Bayesian Information Criterion) (Akaike, 1980) which is recently used in some fields of inverse problems. However, the computation of ABIC is almost impossible in such large problems as our tomographic inversion, since it requires trial and error procedure including calculation of the determinant of the Hessian matrix  $A^T A$ .

An alternative method we applied is the cross validation, which was first proposed by Stone (1974). The principle of the method is very simple. Suppose we have a data set  $d$  sampled from the population. To be precise, first we exclude a datum  $d_i$  from  $d$  and estimate a model  $m = m^{-i}$  from the remainder of the data using a trial value of the roughness parameter, say  $\lambda$ . Then we predict the missing data  $d_i$  from  $m_\lambda^{-i}$  and have the prediction error  $\epsilon_\lambda^{-i}$ . Given all the prediction errors  $\epsilon_\lambda^{-i}$ ,  $i = 1, \dots, N$  by this procedure, where  $N$  is the number of the data in  $d$ , an optimum value of the parameter  $\lambda$  is chosen to be a minimizer of the cross validation score

$$\begin{aligned} f(\lambda) &= \|\epsilon_\lambda\|^2 \\ &= \sum_{i=1}^N [d_i - (g_i, m_\lambda^{-i})]^2 \end{aligned} \quad (21)$$

where  $g_i$  is the  $i$ -th row vector of the data kernel  $G$ . The generalized cross validation given above (Craven and Wahba, 1979) requires large computational efforts. However, if we have a vast amount of data, the cross validation score can be simplified as follows.

Instead of extracting every single datum, we leave out, say, a data set  $d^1$  which is a fraction of the total data set  $d$  through its random sampling. Then we predict  $d^1$  by the model  $m_\lambda^{-1}$  which is estimated from the rest of the data  $d^{-1}$  using a trial  $\lambda$ . The approximate cross validation score is given as

$$f(\lambda) = \|d^1 - G m_\lambda^{-1}\|^2 \quad (22)$$

We find an optimum parameter  $\lambda$  minimizing  $f(\lambda)$  by a trial and error method.

In our problem, we have two parameters  $\sigma_h$  and  $\sigma_v$  to be chosen as a minimizer. It is for this reason necessary to calculate cross validation score  $f(\sigma_h, \sigma_v)$  at many grid points to find the point at which  $f$  is minimal. This is still a time consuming process but an unavoidable one.

### 2.4. Resolution and variance

Since the work of Backus and Gilbert (1968), geophysicists have commonly recognized that a solution of an inverse problem must be consulted with its resolving kernel as well as its a posteriori variance. The result is less valuable unless its reliability is evaluated properly using these quantities. A formulation of the resolution and the covariance matrices in discrete inverse problems are now widely known (Jackson, 1979; Aki and Richards, 1980), which is adopted in tomographic inversions such as Aki et al. (1977) and Dziewon-ski (1984). In our problem, however, the calculation is not straightforward because of the large matrix size. Until recently, the calculation of resolution and variance in the tomographic inversion has been thought to be of great difficulty when the number of parameters exceed several thousand. However, the fact is not so; we have an expedient technique as shown below. Its basic idea is not original, partly given in Humphreys and Clayton (1988) and a simple application is demonstrated in Grand (1987) and Hager and Clayton (1989).

#### 2.4.1. Point resolution and covariance

The a posteriori covariance matrix of the solution of eqn. (15) is given as

$$\begin{aligned} C_m &= (A^T A)^{-1} \\ &= (G^T C_d^{-1} G + D^T C_m^{-1} D)^{-1} \end{aligned} \quad (23)$$

The resolution matrix is then

$$R = C_m G^T C_d^{-1} G \quad (24)$$

Unlike the resolution matrix in the normal damped least squares,  $R$  given above is not a symmetric matrix because  $G^T G$  and  $D^T D$  are not commutative.

The calculation of the inverse matrix in eqn. (23) is impractical in our problem as mentioned previously. However, it is possible to have a column (and row) vector  $c_j$  in  $C_m$  by solving

$$A^T A x_j = e_j \quad (25)$$

where  $e_j$  is the unit vector whose  $j$ -th element is unity. Equation (25) can be solved by the conjugate gradient method as

$$r_0 = e - A^T A x_0$$

$$p_0 = r_0$$

Repeat for  $k = 0, 1, 2, \dots$

$$\alpha_k = (r_k, r_k) / (p_k, A^T A p_k) \quad (26)$$

$$x_{k+1} = x_k + \alpha_k p_k$$

$$r_{k+1} = r_k - \alpha_k A^T A p_k$$

$$\beta_k = (r_{k+1}, r_{k+1}) / (r_k, r_k)$$

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

The term  $A^T A p$  can be calculated by multiplications with row action  $q = A p$  followed by  $A^T q$ . The row vector  $y_i$  of  $R$ , which is called the resolving kernel, for the target point at block  $j$  is given by

$$y_j = x_j G^T C_d^{-1} G \quad (27)$$

The RHS of eqn. (27) can be calculated by a row action at one time on the weighted data kernel  $G' = C_d^{-1/2} G$  stored on a disk file.

The point spread function, i.e., the column vector of  $R$  can be obtained by other means (Humphreys and Clayton, 1988). First we calculate the synthetic data set  $d_j$  given for the model of a unit vector  $e_j$ , in which only the  $j$ -th block

has the unit value. Next the data  $d_j$  is backprojected into the mantle by eqn. (15). The solution is formally given by

$$\hat{e}_j = (G^T C_d^{-1} G + D^T C_m^{-1} D)^{-1} G^T C_d^{-1} d_j \quad (28)$$

Namely

$$\hat{e}_j = R e_j \quad (29)$$

thus,  $\hat{e}_j$  is equivalent to the  $j$ -th column vector in the resolution matrix  $R$ .

The above method is very useful to obtain a covariance, a resolving kernel, and a point spread function when we examine the solution with special attention to a certain target point. However, the effort to calculate them for only one target costs the same as the effort to calculate the solution itself. Global mapping of the reliability of the solution by this method is thus impractical.

#### 2.4.2. Checkerboard resolution

By the concept of resolution we wish to know how the true Earth is reconstructed in the calculated image. A straightforward method is to reconstruct a test pattern from the synthetic data. The method is applied in Grand (1987). For simplicity, we recommend a checkerboard pattern in both the horizontal and vertical coordinates. First we construct a model of checkerboard pattern as

$$m_{cb} = \sum_j c_j e_j \quad (30)$$

where  $c_j$  are the coefficients 1 or  $-1$  defining the pattern. Next, the synthetic data set is generated using the data kernel for the real data set  $G$  as

$$d_{syn} = G m_{cb} \quad (31)$$

The  $d_{syn}$  is then backprojected into the mantle to reconstruct the test pattern. The reconstructed image can be written as

$$\hat{m}_{cb} = R \sum_j c_j e_j = \sum_j c_j \hat{e}_j \quad (32)$$

which is a superposition of the point spread functions  $\hat{e}_j$ . Although the equations above are trivial, they clearly indicate the role of the checkerboard resolution. The method is recommended in its application for several checkerboard patterns with different wavelengths. If one uses a robust esti-

mate in backprojection, i.e., introducing a residual dependent weight on the data, the amplitude of the test pattern must be about the same as that of the true Earth.

#### 2.4.3. Global variance

The equation for the covariance matrix given in eqn. (23) shows the propagation rule from the a priori errors on the data and the smoothness defined by  $C_d$  and  $C_m$ , respectively, to the a posteriori error on the solution defined by  $C_{\hat{m}}$  through the backprojection. Instead of calculating the RHS of eqn. (23), we use an alternative method to estimate  $C_{\hat{m}}$ , which is to simulate error propagation from gaussian noise input. We replace the observed data  $d$  and the a priori smoothness  $0$  in the RHS of eqn. (11a) by gaussian noise with zero mean, respectively. The covariance matrix of the input noise must be the same as that used in the real data inversion. The solution of the resultant inversion is a projection of the input noise into the model space. Having  $N$  such solutions  $\hat{m}^n$ , ( $n = 1, \dots, N$ ) given from  $N$  sets of independent noise, we can obtain an element  $\hat{c}_{ij}$  of the approximate a posteriori covariance matrix as

$$\hat{c}_{ij} = \frac{1}{N} \sum_{n=1}^N m_i^n m_j^n \quad (33)$$

The corresponding approximate resolution matrix is formally given from eqn. (24). However, it is practically impossible to calculate all the elements in  $R$  in our problem.

### 3. Application

The method of global tomographic inversion described above has been applied to the real data. In the following subsections, we present the description of the model and the data, optimization of the roughness parameters, and the summary of the final solution consulting with maps of resolution and variance.

#### 3.1. Model and data

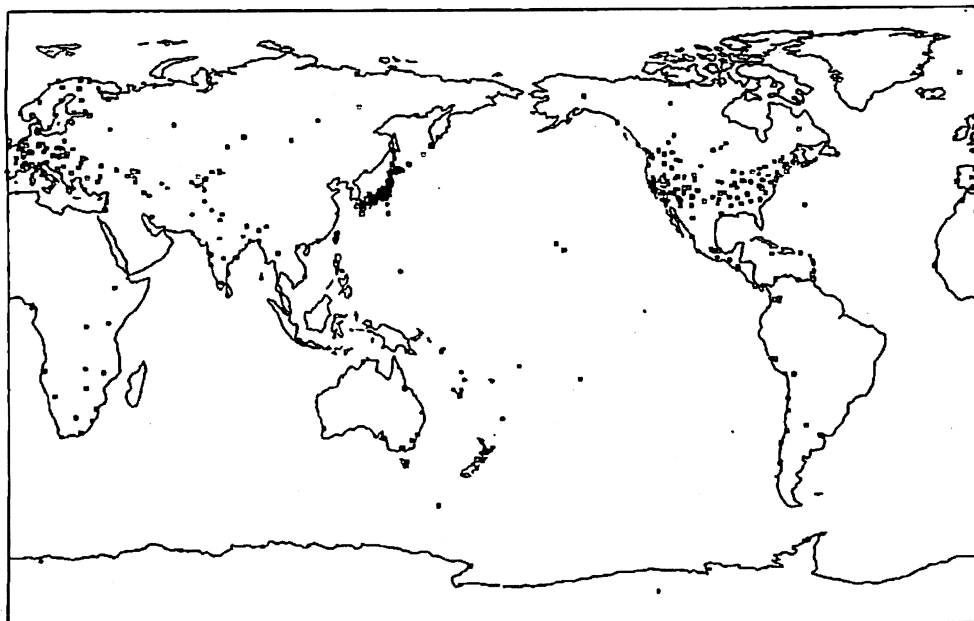
The spherically symmetric model used as the starting model is based on Herrin et al. (1968); the

crust and the uppermost mantle are smoothed and no discontinuities are included. The velocities are given at 64 nodes which distribute according to the formula similar to eqn. (1). The travel time table is constructed by interpolating these velocities using the function  $v(r) = ar^b$  in each layer. The slowness perturbation is modeled into 32,768 conical blocks as mentioned previously (Fig. 1). A coarse model with 4,096 ( $16 \times 32 \times 8$ ) blocks is, however, used in the first iteration to help fast convergence.

The travel time data are extracted from the bulletin of the ISC for the period of 21 y from 1964 to 1985 excluding 1970. Distributions of the stations in 1964 and 1985 are shown in Fig. 3. The total number of all the events stored on the original magnetic tapes is about 330,000. First, we selected the events such that  $m_b$  was 4.0 or greater and the number of readings was 20 or greater. This selection reduced the number of events to about 75,000; however, it is still impractical to invert them all at once. In the case of the slowness model with 32,768 blocks, the data kernel  $G$  for up to about 2000 events is allowed to be kept on the memory of our computer. Therefore, we extracted ten data sets containing about 2000 events for each and inverted them to produce ten models. The final model is obtained as weighted average of these ten models and an approximate variance of the solution is estimated from the variance among them.

The 2000 events in each data set are extracted from the 75,000 events as follows. The 21 y from 1964 to 1985 are divided into ten periods of four years each overlapping two y, such as '77-'81, '79-'83, '81-'85. In each period, we selected the events so as to be distributed as uniformly as possible in both horizontal and vertical extents. The process is as follows. The upper mantle (0–800 km) is divided into  $32 \times 64 \times 8$  cells in latitude, longitude, and depth respectively. The surface cell size is  $5.625^\circ \times 5.625^\circ$  and the depth boundaries are 33 km, 100 km, 200 km, 300 km, 400 km, 500 km and 600 km. Sorting the events in each cell by the number of readings with weighted counts of 1:2:4 for  $eP:P:iP$ , we selected up to three events in a cell. The number of events in each cell is limited to one per month, in order to avoid a

6401 NST= 586



8512 NST=1466

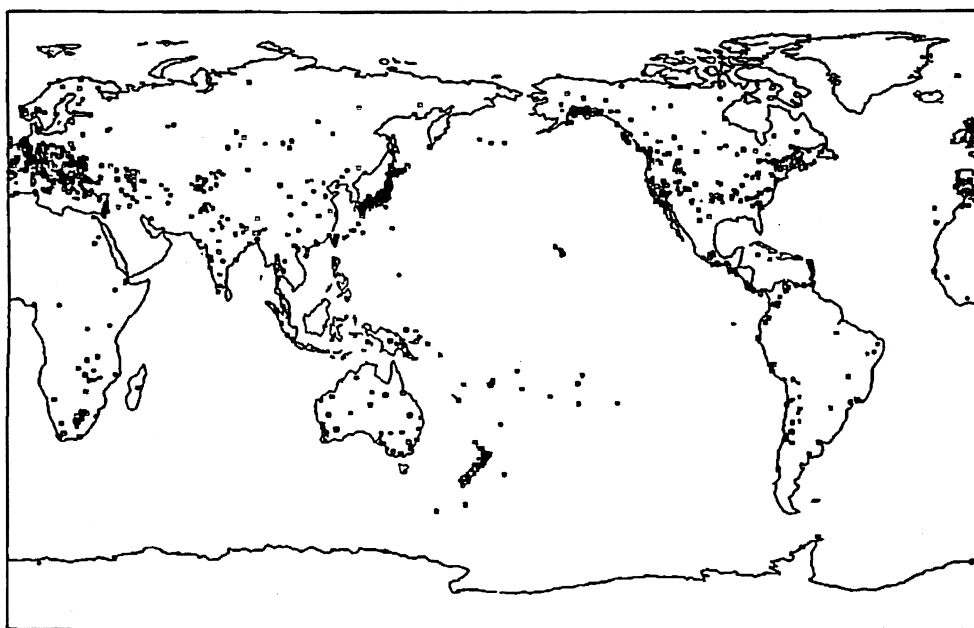


Fig. 3. Distribution of seismograph stations in January, 1964 (upper) and December, 1985 (lower).

concentration due to aftershocks. As a result, 22 thousand events have remained totally in the ten data sets, for each of which the number of events

and readings are tabulated in Table 1. Figure 4 shows the distribution of the events in the four depth ranges. One fourth of them are doubly used

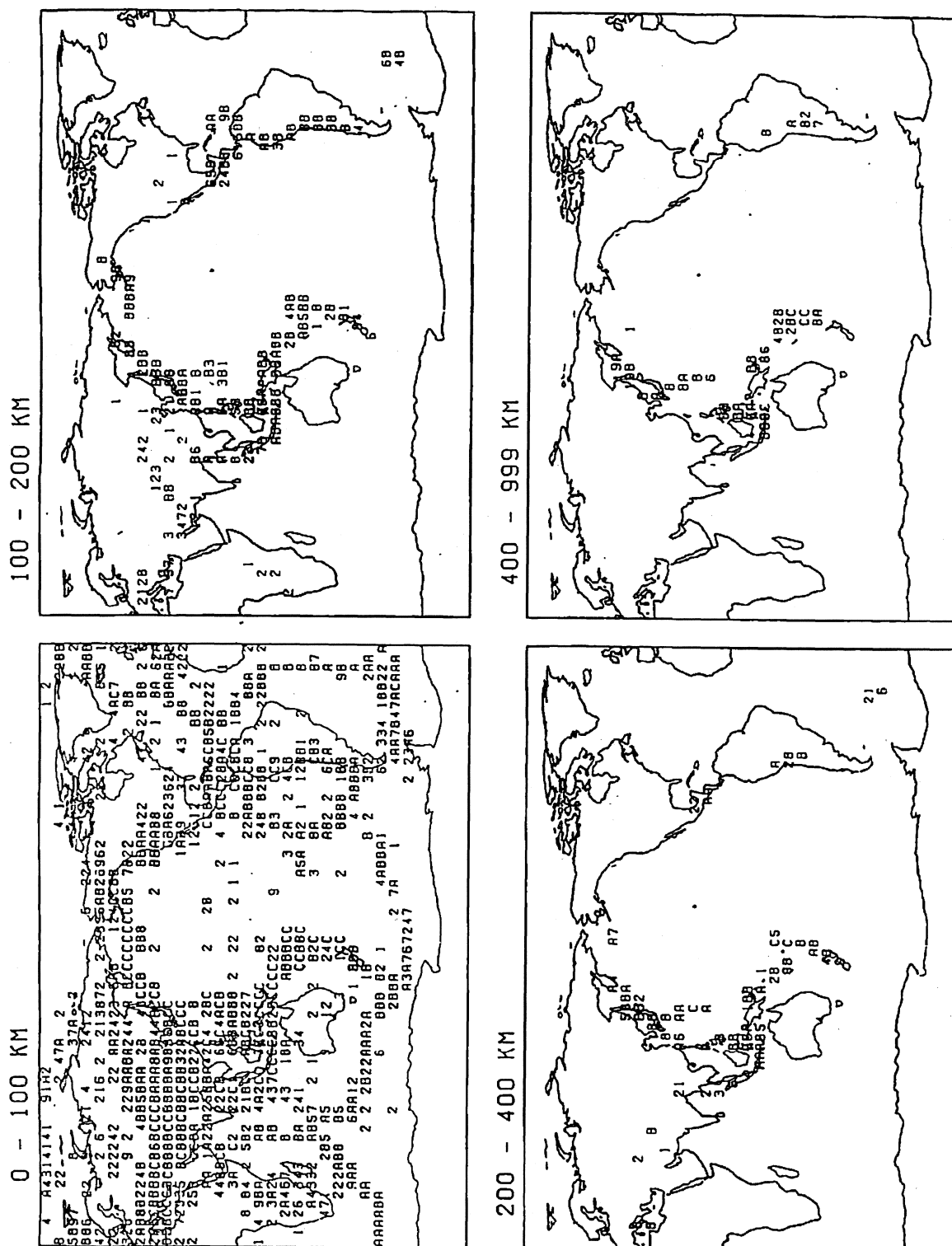


Fig. 4. Distribution of events used in this study. The number of events in a  $5.625^\circ \times 5.625^\circ$  cell is given by numerals or alphabets. (Numerals  $< 10 \leq A < 20 \leq B < 50 \leq C < 100$ ).



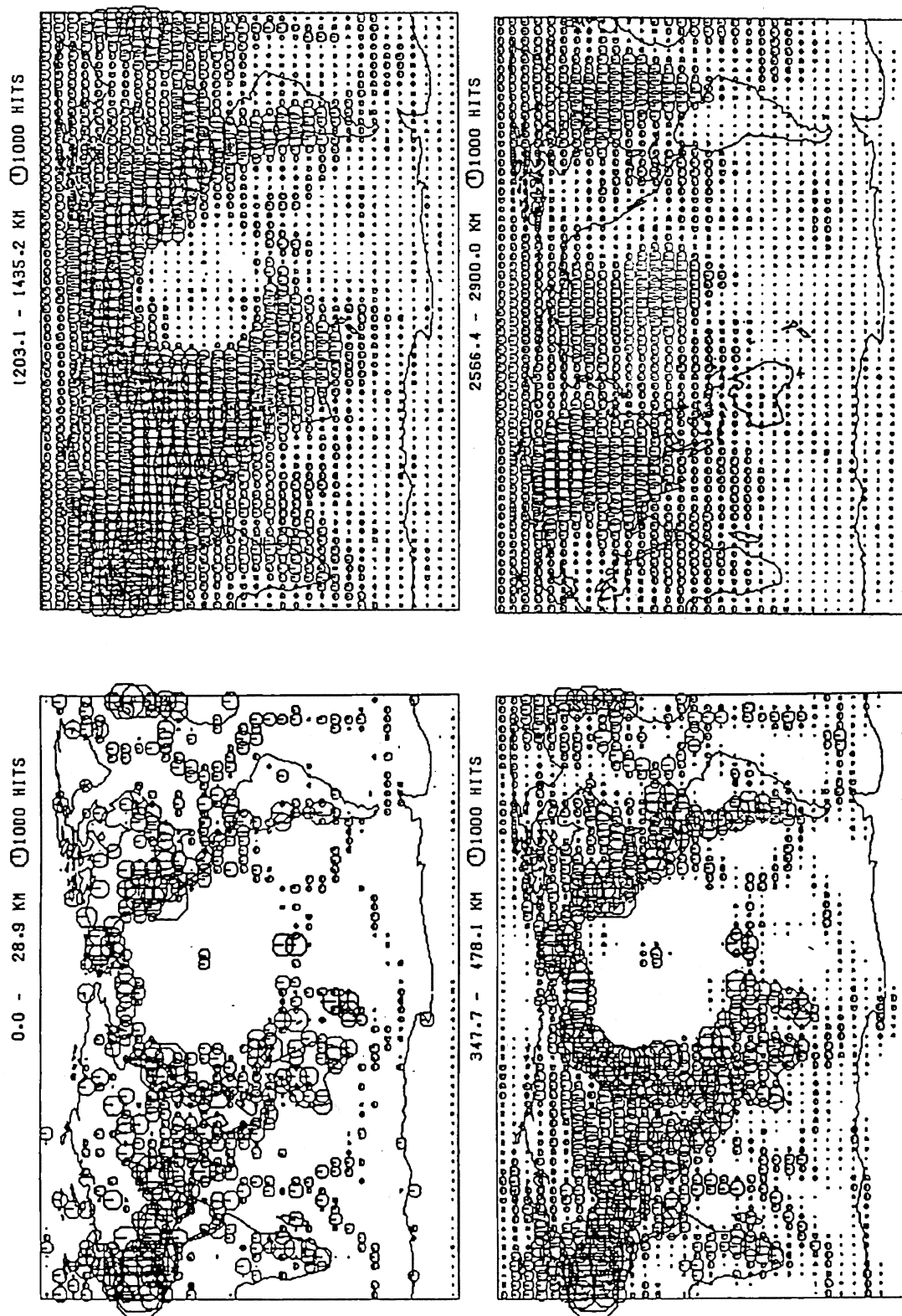


Fig. 5. Hit count map for the data set No. 9 in Table 1. The number of the rays in a cell is expressed by different size of circles.

TABLE 1  
List of data sets

No.	Year	$N_{ev}^a$	$N_p^b$
1	64-67	2,044	154,691
2	66-69	2,164	184,508
3	68-71	1,949	176,336
4	70-73	1,972	174,164
5	72-75	2,215	231,100
6	74-77	2,260	247,907
7	76-79	2,279	266,200
8	78-81	2,272	275,826
9	80-83	2,292	278,610
10	82-85	2,337	316,690

<sup>a</sup> Number of events; <sup>b</sup> number of readings.

due to the overlap of the periods, which are mainly located under the ocean bottom. The total number of the paths are 2.3 million and the average number of observations for an event is 105. All the travel times are then corrected for ellipticity (Dziewonski and Gilbert, 1976) and elevation of the stations. The phases at epicentral distances  $> 95^\circ$  in the ISC bulletin were omitted. Figure 5 shows the hit count map, which plots the number of rays passing through each block. This gives information on the reliability of the solution as well as the resolution and variance maps shown later.

Before proceeding to the next subsection, we briefly describe a computational aspect of our inversion process. The calculations have been performed on the main frame computer FACOM M780 (20 MFLOPS max.). All the calculations of solution, cross validation, checkerboard resolution and variance involve backprojection process. The approximate size of the data kernel  $G$  for 2000 events is 100 megabytes, a maximum size of the memory on our computer. It takes 5 min in CPU time to repeat 64 iterations in CG method for a backprojection. The construction of a data kernel  $G$  is comparable in CPU time. The backprojections are performed keeping  $G$  on the memory, because reading it from a disk file takes more than ten times of the CPU time.

### 3.2. Cross validation

Prior to inverting all the data, we determined the optimum values of the roughness parameters

$\sigma_h$  and  $\sigma_v$  in eqn. (9) by a simplified technique of cross validation described in Section 2.3.3. The data set used for the cross validation contains 2034 events and 259,143 paths, which is similar in its size to a data set shown in Table 1. First, all the events are relocated with the spherically symmetric starting model. Next we divided the data set to  $d_1$  and the remaining  $d_2$ ; the latter is inverted to trial solutions and the former is predicted. If data set  $d_1$  is taken to be too small, the result tends to depend on a particular choice of  $d_1$  and the estimated optimal values become unreliable. If the data set  $d_2$  is taken to be too small, on the other hand, the results tends to depend on the size of  $d_2$  and the estimated optimal values become biased. Taking into account the trade-off, we took one eighth of the  $d$  as  $d_1$ . We cut out one reading in every eight readings in chronological order. The  $d_1$  of 32,392 readings were thus left out

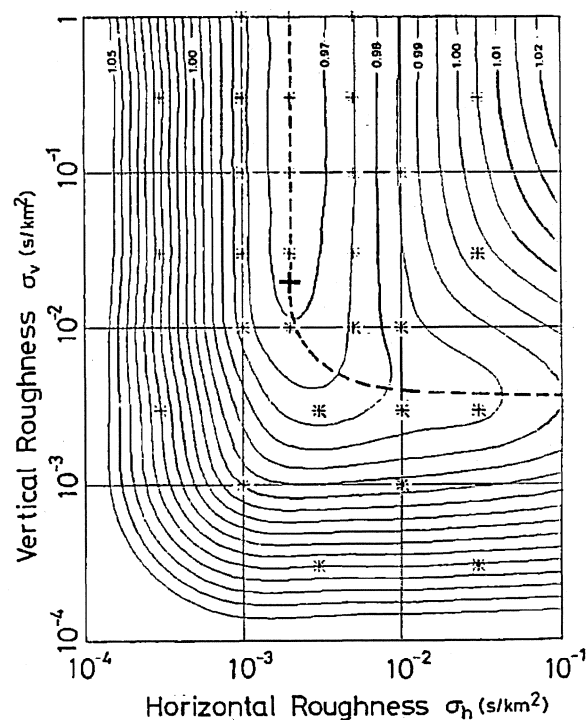


Fig. 6. Contour map of the cross validation score for the data set similar to No. 9. The prediction errors are measured at the points indicated by asterisks. The optimum pair of parameters  $\sigma_h$  and  $\sigma_v$  adopted in the inversion is shown by large plus. Dashed curve indicates the trough along the prediction error minima. Contour interval of the prediction errors 0.005 s.

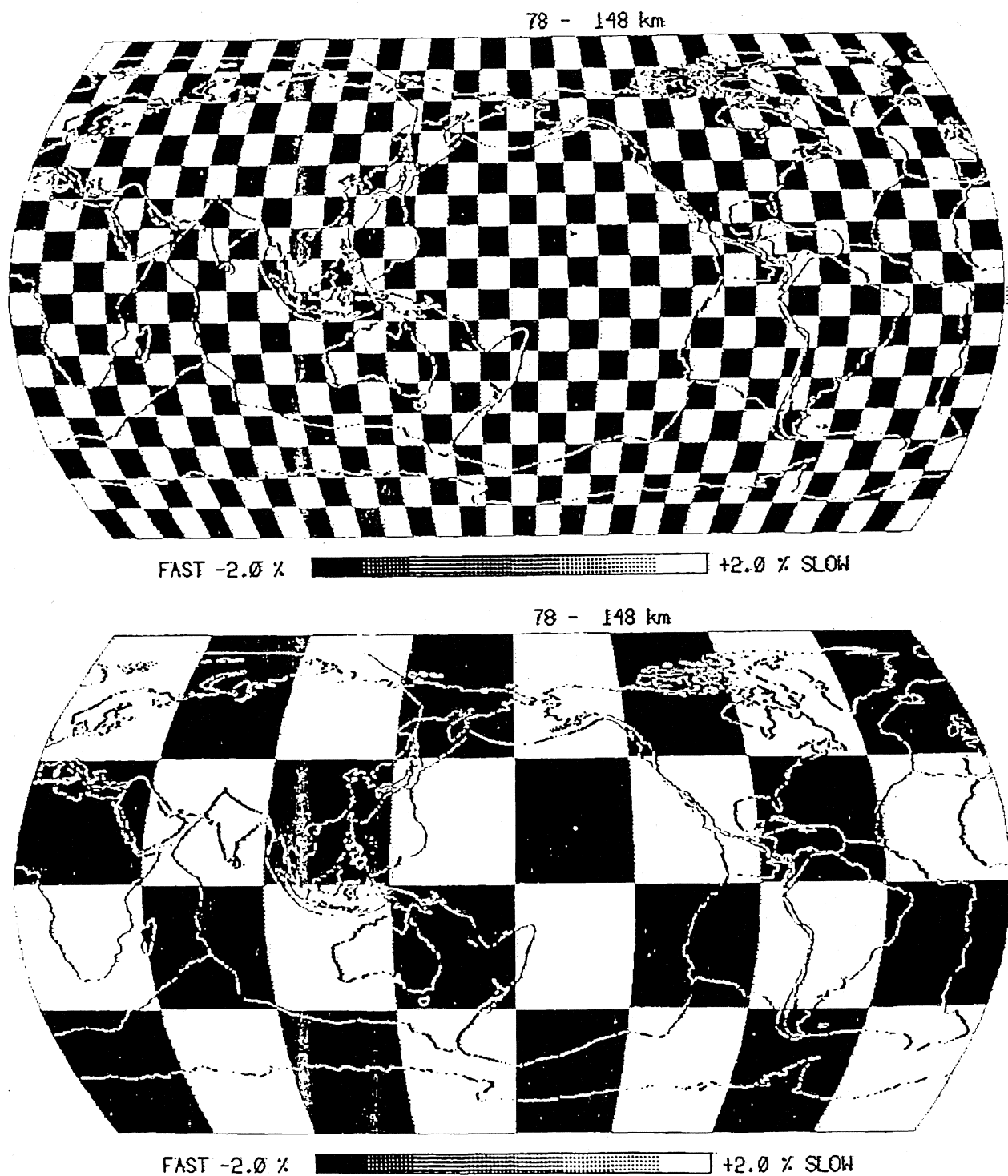


Fig. 7. Horizontal sections of test patterns for the checkerboard resolution. The alternate patterns of  $\pm 2\%$  anomaly in slowness extend both horizontally and vertically.

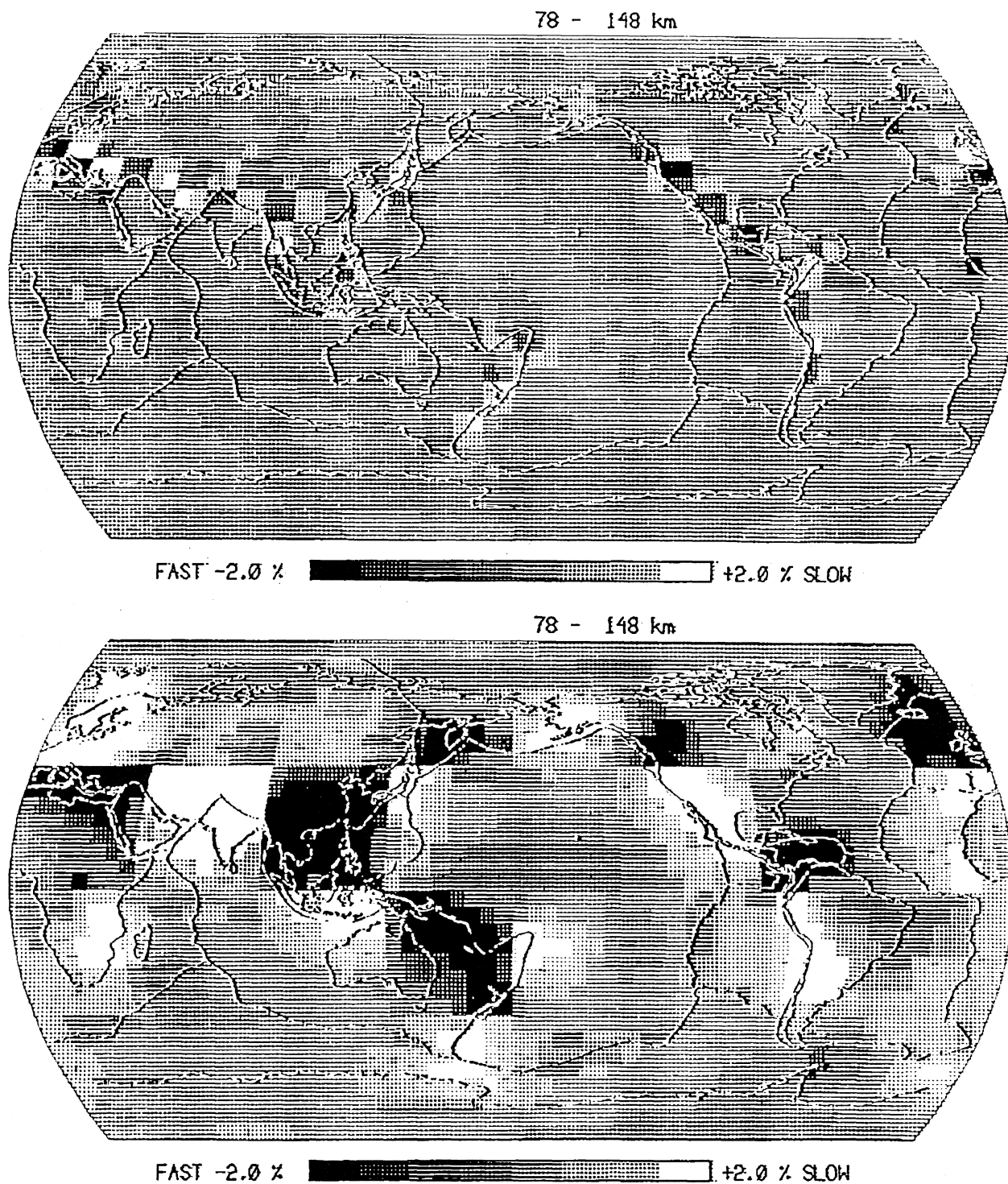


Fig. 8. Checkerboard resolution maps at four different depths calculated for the data set No. 9. Complete resolution should reproduce the patterns shown in Fig. 7.

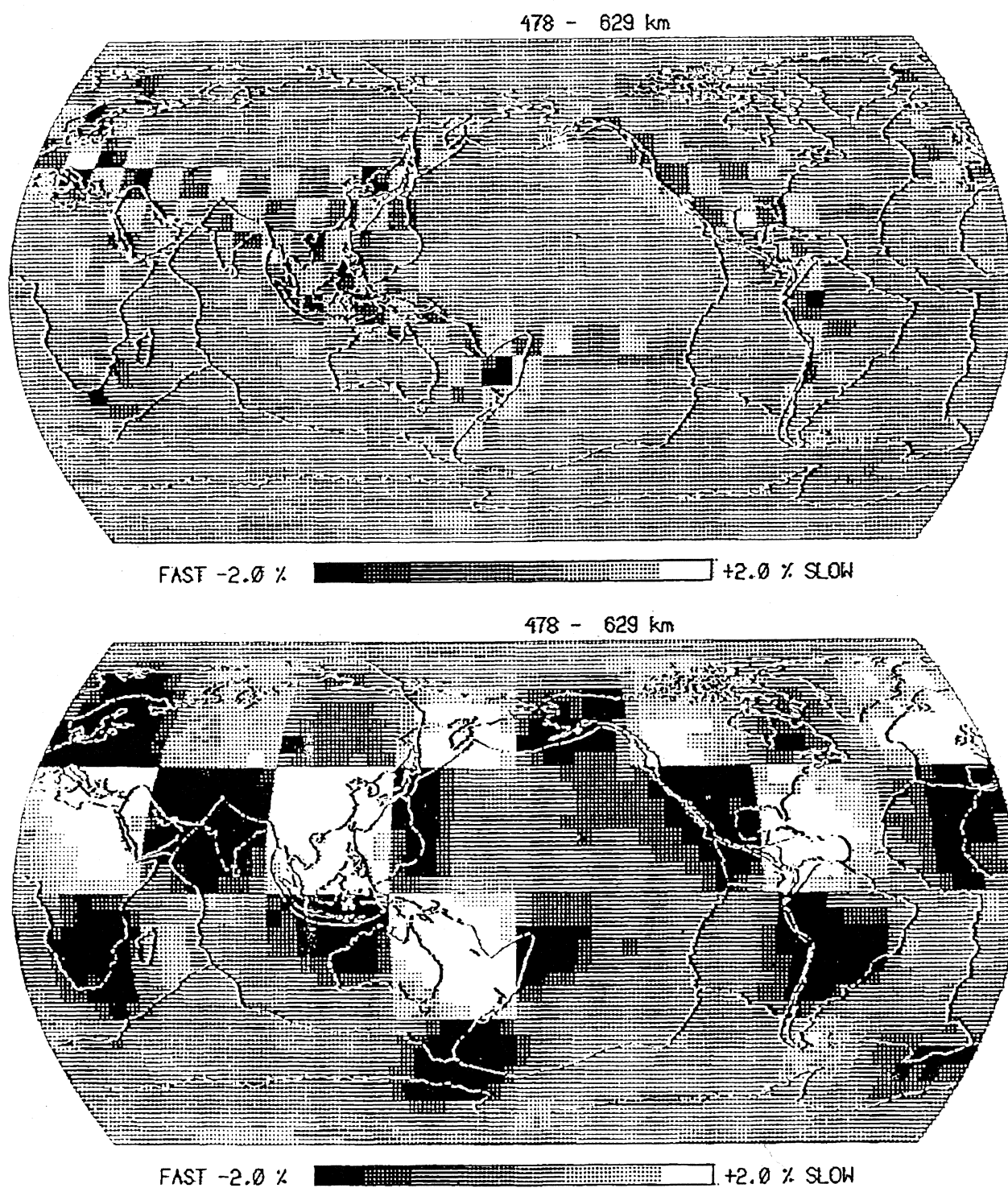


Fig. 8 (continued).

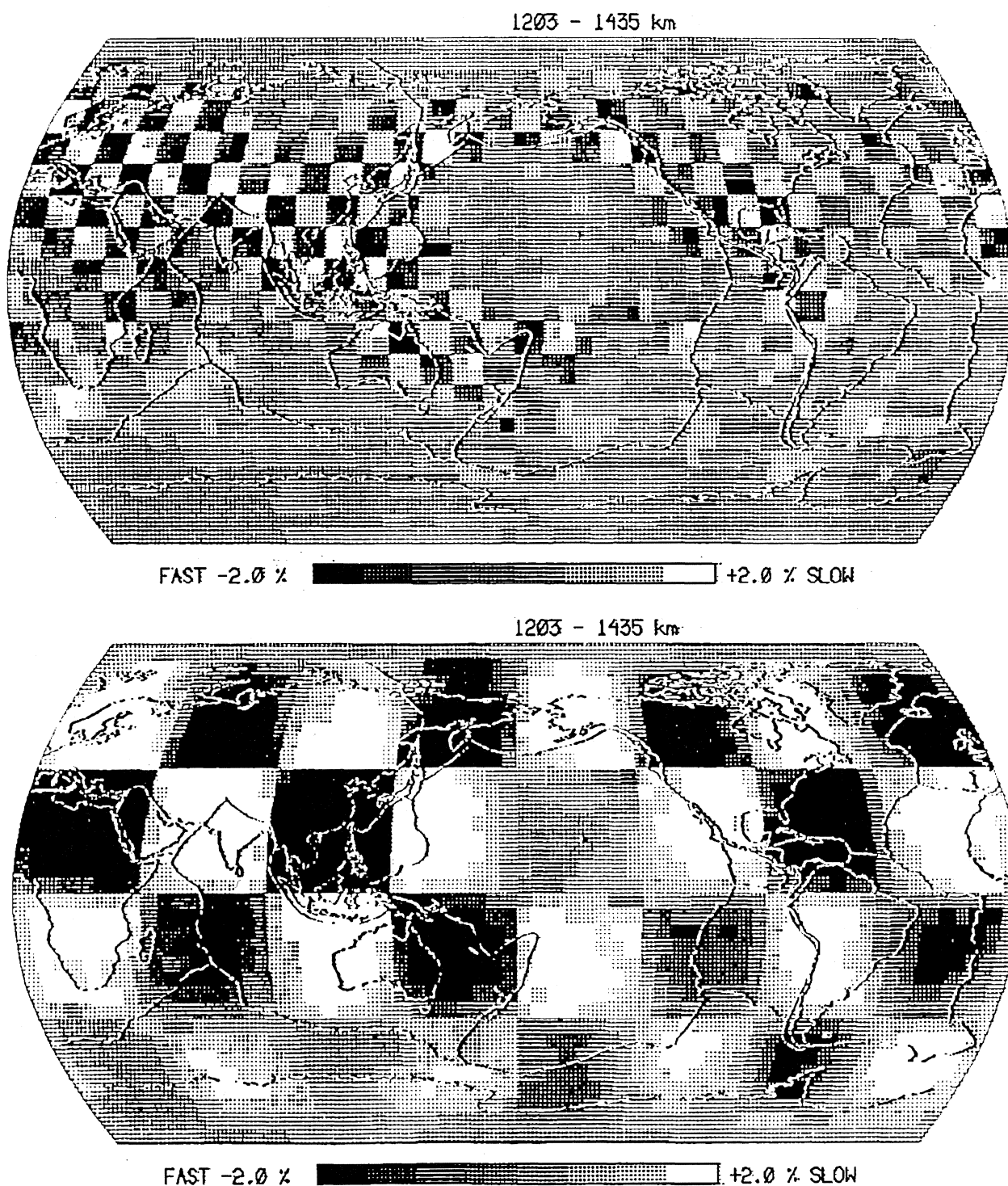


Fig. 8 (continued).



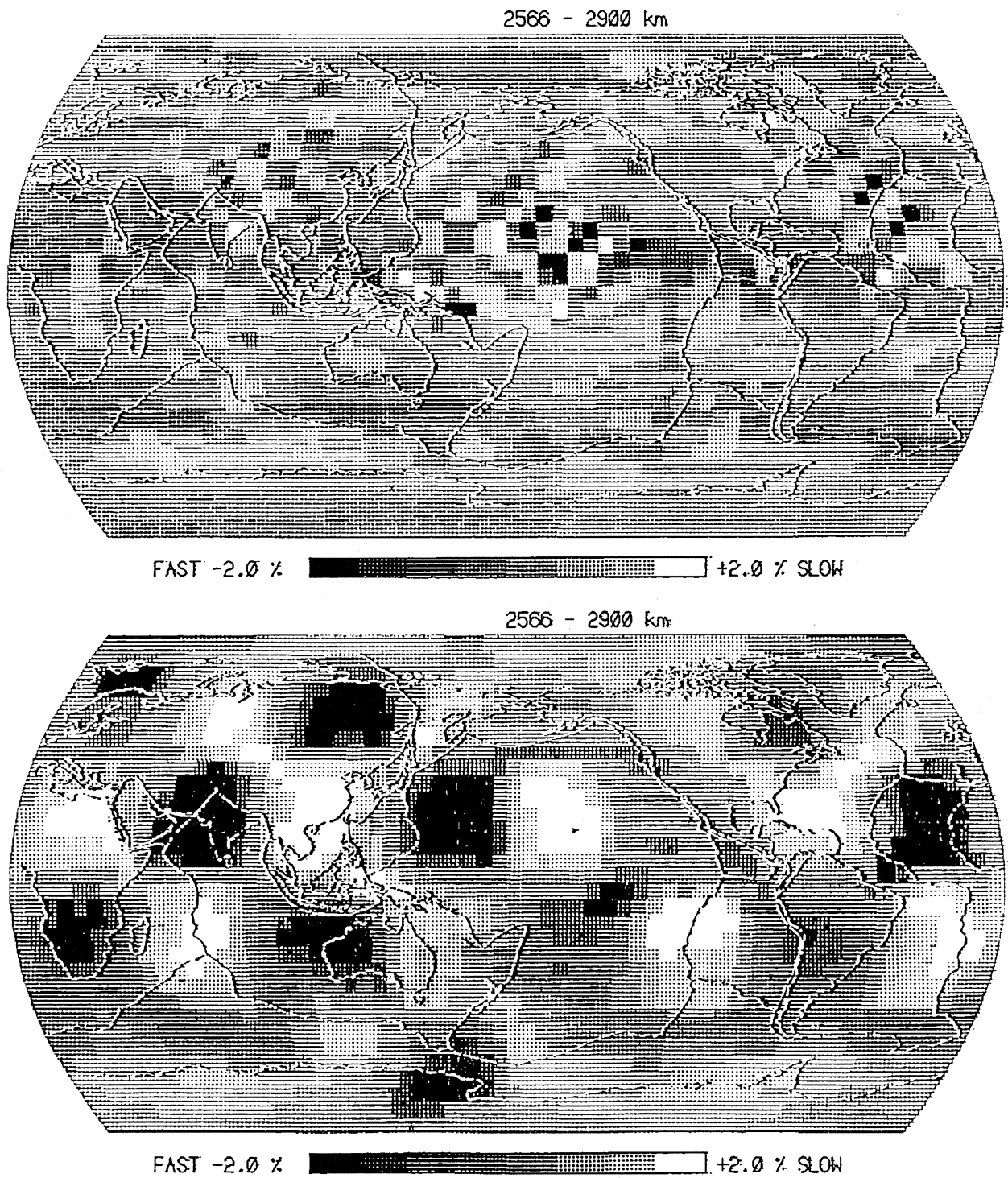


Fig. 8 (continued).

from the  $d$  of 259,143 readings. The seven eighths of the data  $d_2$  are inverted to obtain trial solutions with various roughness parameters. The back-projection was performed with the CG method as described in the last section. The iterations were stopped at 16 to save CPU time, which is sufficient to converge for the blocks well covered with the paths. We performed 64, rather than 16, iterations in the final inversion described in Section 3.4.

Figure 6 shows the result of cross validation. Trial solutions are calculated for thirty different pairs of  $\sigma_h$  and  $\sigma_v$ , which are shown by small symbols. The corresponding r.m.s. prediction errors for the solutions are then interpolated and shown by contour lines at 0.005 s intervals. The horizontal roughness can be well constrained from this contour map as one giving the least prediction error. For the vertical roughness, on the other hand, only the lower bound can be set from this map: prediction error does not increase even if the vertical damping of smoothness is relaxed. This under determined value suggests that the vertical division into the 64 nodes might have been too coarse to follow the true vertical heterogeneity of the mantle. We use, in this paper, the smallest vertical roughness among those yielding minimal prediction errors. The adopted values of the optimum parameters are  $\sigma_h = 2 \times 10^{-3}$  and  $\sigma_v = 2 \times 10^{-2}$  (s km<sup>-2</sup>) as shown by the large cross in Fig. 6.

### 3.3. Resolution

Prior to inverting all the data, we examine the checkerboard resolution. Figure 7 shows a horizontal section of the test pattern. Two sizes of the pattern are tested in order to see the resolutions in different wavelengths. A black or a white area involves  $2 \times 2 \times 1$  blocks (top), and  $8 \times 8 \times 4$  blocks (bottom) in latitude  $\times$  longitude  $\times$  depth. The test patterns consist of slowness perturbations of  $\pm 2\%$ , which are plausible values for the true perturbation in the mantle.

The synthetic travel times are calculated for the ray paths in data set No. 9 and they are back-projected into the mantle. The roughness parameters, the number of iterations, and the weight on

each datum used are the same values as in the real data inversion shown later. Although the resolution should be calculated for all the ten data sets, we have calculated it for only the data set No. 9 to save CPU time. Therefore, the resolution of our final model should be somewhat better than that calculated here.

Figure 8 (a–d) shows the reconstructed image of the test patterns. In the depth range 78–148 km, the small scale resolution (top) is good only at the limited areas, around the Pacific, South Asia and Europe. The areas along the mid oceanic ridges are not resolved at all, though there are some rays passing through them (Fig. 5). The large scale resolution (bottom) is better than the small scale. However, the ocean and the high latitude area are still not resolved. In the depth range 478–629 km, the resolution increases. The large scale pattern is well reconstructed except for the mid oceanic and polar regions. In the depth range 1203–1435 km, the original large scale pattern is reconstructed quite well. Just above the CMB (2566–2900 km), the central Pacific and the central Atlantic are well resolved. The resolution on the average is, however, slightly poorer than that in the layer above it (2253–2566 km). Possible reasons for this are that (1) the number of rays decreases (see Fig. 5), and that (2) the horizontal smoothness of the solution effectively increases, as the vertical smoothness is relaxed at the CMB.

Resolution in polar regions is poor in all the depths. This is due to the effect of smoothing and slow convergence, because block sizes in longitude are very small in polar regions.

### 3.4. Inversion of the data

The inversion is performed by the algorithm described at the beginning of the last section, i.e., the relocation, backprojection and the refinement of the spherically symmetric model are repeated until the solution converges. The optimum roughness parameters  $\sigma_h = 2 \times 10^{-3}$  and  $\sigma_v = 2 \times 10^{-2}$  determined by the cross validation are used only at the final backprojection. During the preceding iteration processes we adopted the smaller values  $\sigma_h = 1 \times 10^{-3}$  and  $\sigma_v = 1 \times 10^{-2}$  to suppress roughness more effectively to give a stable starting



model for the final iteration. In each iteration, we have ten solutions of slowness perturbation corresponding to the ten data sets shown in Table 1. The ten solutions are averaged with weights proportional to the number of observations contained in each data set. Then the spherically symmetric model is revised by averaging perturbations in the shells, where the contributions from blocks containing no ray are excluded. Although the ray coverages are not uniform among the ten data sets due to the difference in station distribution, the anomaly patterns and their amplitudes of the ten individual solutions are quite similar.

### 3.4.1. 1-D model

The starting and the final models of the spherically symmetric Vp structure are listed in Table 2. They are plotted in Fig. 9, as well as the original model of Herrin et al. (1968). As the result of inversion the crust and uppermost mantle are changed but the lower mantle is little changed. From the surface to a depth of 100 km, the velocity increases by about  $0.05 \text{ km s}^{-1}$ . The largest change of  $0.07 \text{ km s}^{-1}$  occurred at the depth of 50–70 km and, intriguingly, a low-velocity layer (LVL) emerged directly below it. The LVL seen in Fig. 9 is small. A more distinct LVL

TABLE 2  
The starting and the final spherically symmetric Vp model

Node No.	Depth (km)	Starting Vp model ( $\text{km s}^{-1}$ )	Final Vp model ( $\text{km s}^{-1}$ )	Node No.	Depth (km)	Starting Vp model ( $\text{km s}^{-1}$ )	Final Vp model ( $\text{km s}^{-1}$ )
0	0.000	6.000	6.051	33	845.949	11.165	11.165
1	5.324	6.053	6.104	34	893.164	11.252	11.252
2	11.914	6.157	6.211	35	941.652	11.341	11.341
3	19.777	6.393	6.451	36	991.406	11.427	11.425
4	28.906	6.934	6.993	37	1042.434	11.512	11.510
5	39.309	7.652	7.715	38	1094.727	11.596	11.593
6	50.977	8.065	8.133	39	1148.293	11.681	11.678
7	63.918	8.079	8.149	40	1203.125	11.768	11.762
8	78.125	8.094	8.137	41	1259.230	11.853	11.844
9	93.605	8.110	8.126	42	1316.602	11.936	11.927
10	110.352	8.130	8.146	43	1375.246	12.019	12.010
11	128.371	8.158	8.174	44	1435.156	12.101	12.093
12	147.656	8.195	8.194	45	1496.340	12.184	12.176
13	168.215	8.242	8.222	46	1558.789	12.266	12.258
14	190.039	8.299	8.278	47	1622.512	12.344	12.336
15	213.137	8.366	8.345	48	1687.500	12.420	12.418
16	237.500	8.444	8.411	49	1753.762	12.496	12.502
17	263.137	8.534	8.488	50	1821.289	12.572	12.578
18	290.039	8.637	8.591	51	1890.090	12.649	12.655
19	318.215	8.752	8.705	52	1960.156	12.733	12.737
20	347.656	8.880	8.852	53	2031.496	12.818	12.820
21	378.371	9.022	9.017	54	2104.102	12.903	12.904
22	410.352	9.177	9.172	55	2177.980	12.986	12.987
23	443.605	9.345	9.339	56	2253.125	13.074	12.072
24	478.125	9.526	9.513	57	2329.541	13.163	13.157
25	513.918	9.720	9.697	58	2407.227	13.254	13.248
26	550.977	9.924	9.901	59	2486.182	13.352	13.346
27	589.309	10.137	10.114	60	2566.406	13.453	13.445
28	628.906	10.358	10.345	61	2647.900	13.548	13.538
29	669.777	10.582	10.581	62	2730.664	13.619	13.609
30	711.914	10.800	10.800	63	2814.697	13.655	13.645
31	755.324	10.978	10.977	64	2900.000	13.670	13.660
32	800.000	11.077	11.077				

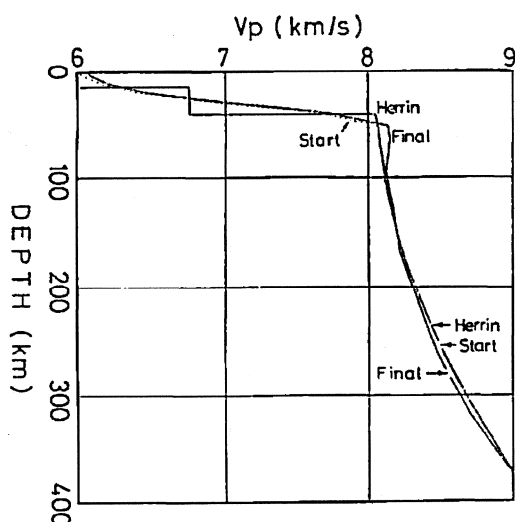


Fig. 9. Spherically symmetric models in the upper mantle. Dashed curve indicates the "smoothed Herrin" model used as a starting model in this study. Two solid curves indicate the original model of Herrin et al. (1968) and the final model given after five iterations of relocation and 3-D backprojection. A low velocity layer of the final model should be noted.

should exist, however, since the LVL resulted at every iteration, which was removed at the next iteration to simplify the calculation of ray path and travel time table. There are no signs for the existence of the 400 km and 650 km discontinuities as is expected for the framework of our inversion.

Figure 10 shows the plot of travel time residuals for the data set No. 9. The residuals after the first relocation using the initial symmetric Earth model is shown at the top, and those after the final 3-D inversion at the bottom. The standard deviation of the readings  $P$  (excluding those  $iP$  and  $eP$ ), to which we assign a standard reading error of 0.3 s, is reduced from 1.43 s (top) to 1.25 s (bottom).

### 3.4.2. 3-D model

The final solution of 3-D slowness perturbation is shown in Fig. 11. The four depth ranges correspond to those of the resolution map (Fig. 8). We summarize the features that emerged from this figure.

In the depth range of 78–148 km, the slowness distribution is strongly correlated to the surface

tectonics (Fig. 11a). The back-arc basins in the western Pacific, from the Bering Sea at the northern end, to the Tasman Sea at the southern end, including those behind the Sunda arc, all exhibit large slow anomalies of about 1.5–2.5%. The Caribbean Sea is also slow, but the Scotia Sea is not resolved. The Phanerozoic orogenic belts, i.e., the western part of the North America, East Asia, Zagros mountains and South Europe, are also slow. The Andes and the Himalayas have no slow anomalies in this depth range. At the depth range 29–78 km, however, there exist large slow anomalies about  $1.5 \times 2.0\%$  beneath those mountains.

The regions along the ridges have been expected to have slow anomalies from surface wave studies (Woodhouse and Dziewonski, 1984; Nakanishi and Anderson, 1984; Suetsugu and Nakanishi, 1987). However, our result shows the slow anomalies only in a limited area; in northernmost Atlantic around the Iceland hot spot, the northernmost East Pacific Rise, the Northern Indian Ocean, the Red Sea, and the Great African Rift Valley, all of which are close to the continents. The expected slow anomalies at the mid oceanic ridges far from the continents are not seen due to small resolving power (see Fig. 8a).

On the other hand, remarkably fast anomalies exist beneath the stable continents: the Canadian shield, Greenland, the Fennoscandian shield, Siberian shield, Indian Peninsula, the African shield, and Australia. The fast anomalies in these areas are about 1–2%. No distinct fast anomalies are seen beneath Brazil and Antarctica due most probably to low resolution. Other remarkable anomalies exist at the ocean side of some major trenches; the Kuril–Japan trench, the Tonga–Kermadec trench and the Java trench. The fast anomalies in these areas reach about 2%. The resolving power at these areas is not so bad.

The anomalies shown above are generally seen in the upper mantle at the depths down to about 350 km; the structural pattern changes around this depth. The layer 478–629 km in the transition zone is shown in Fig. 11b. Most of the back arc basins which exhibit slow anomalies at the depth 78–148 km change to fast anomalies. This change may be due in part to the existence of high velocity slabs. Such change is also seen in the S-wave

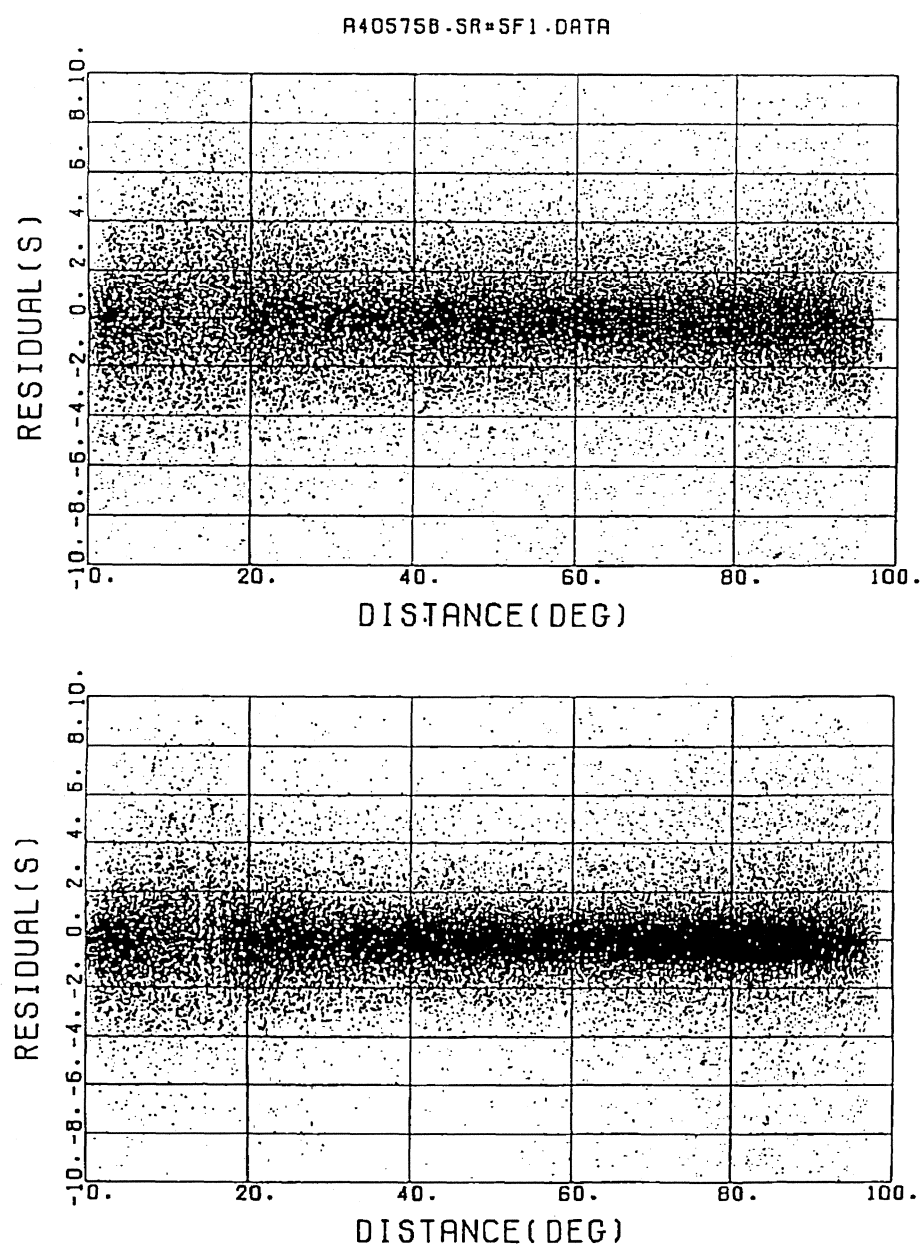


Fig. 10. Plot of travel time residuals against epicentral distance for the data set No. 9. Upper, the initial residuals after the first relocation; lower, the final residuals after the fifth iteration.

structure of Suetsugu and Nakanishi (1987); however, the transition depth is 120–180 km in their model. The orogenic belts of Western America, East Asia, and Europe also change to the fast regions. The remarkable structure around this depth is the widely distributed slow anomalies from the northern Indian Ocean, through Fennos-

candia, to the arctic area. The small scale resolution around these areas is not so good, but large scale anomalies seem to be reliable. There is also a slow anomaly belt from the northeastern Pacific to Antarctica, although the anomalies under the mid Pacific Ocean are less reliable (see Fig. 8b, bottom). We thus find the large scale anomalies in the

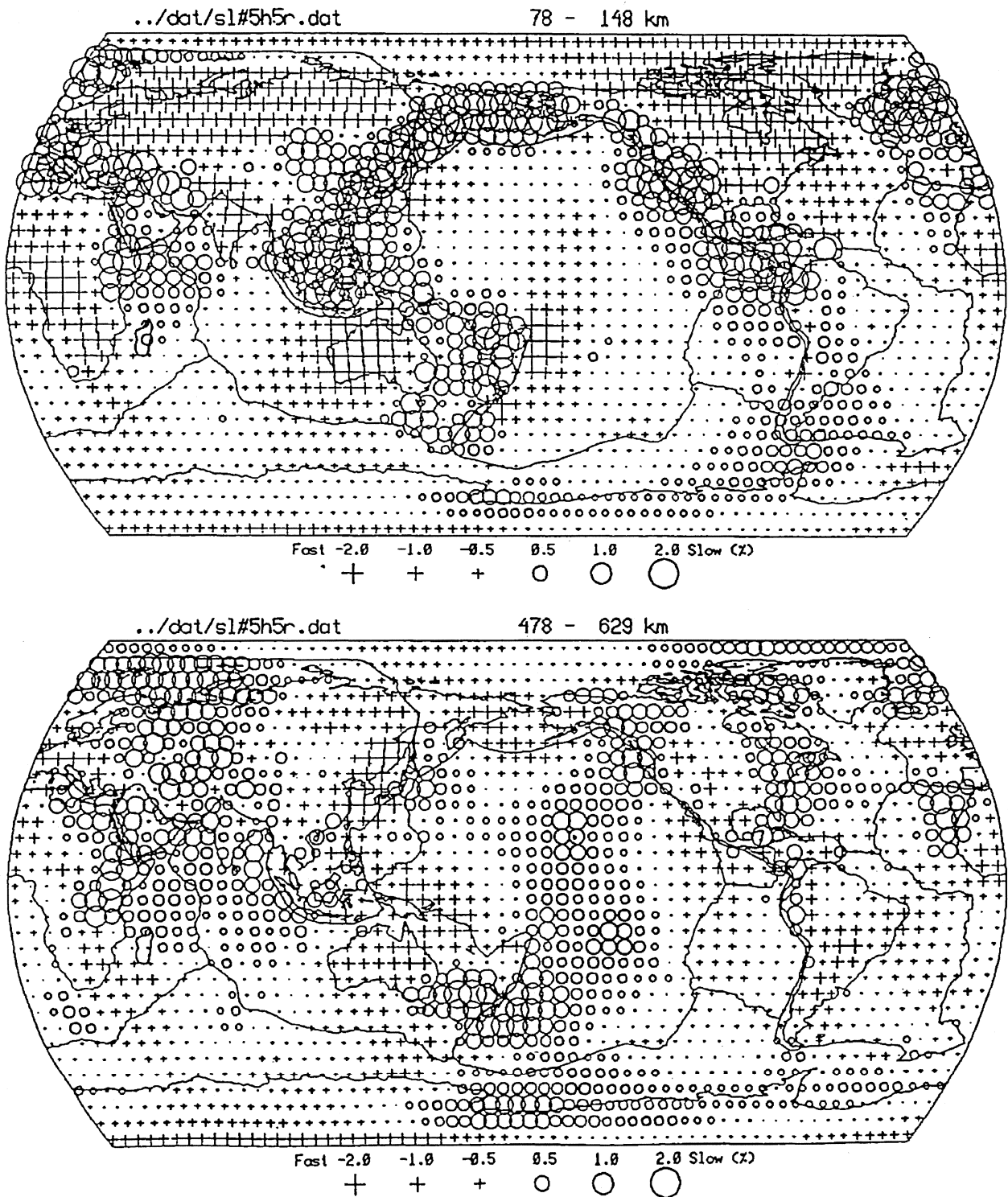


Fig. 11. Slowness anomaly maps at the four different depths. The corresponding resolution maps are given in Fig. 8. Fast and slow anomalies are indicated by crosses and circles, respectively. The size of the symbols is proportional to the square root of the perturbation.

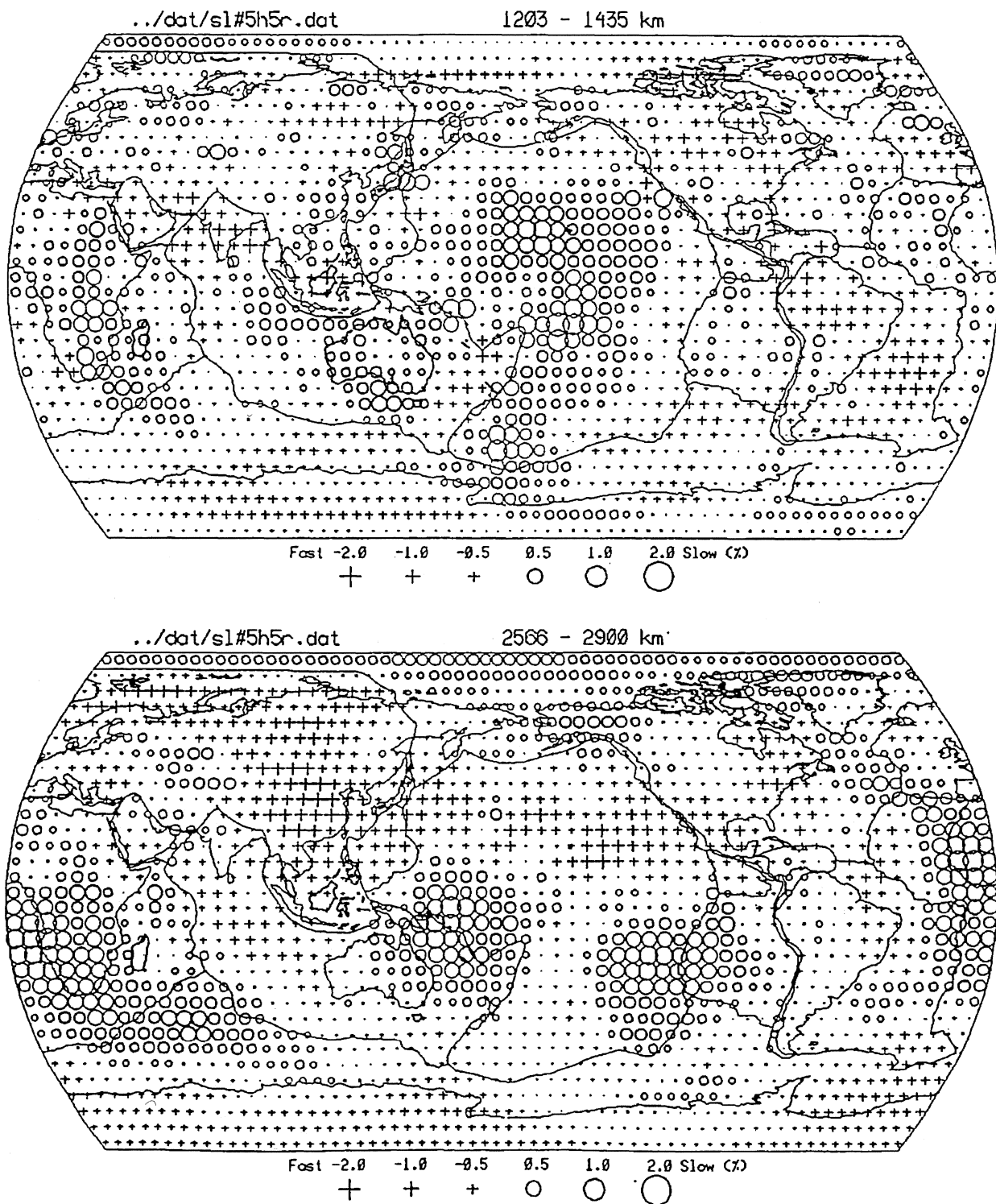


Fig. 11 (continued).

transition layer, i.e., slow at the South Pacific and Western Eurasia, and fast at South America and its antipodal region of the western Pacific. This pattern coincides to the  $l=2$  anomaly given from the free oscillation spectra by Masters et al. (1982), and the surface waves studies such as Nakanishi and Anderson (1984), Woodhouse and Dziewonski (1984) and Tanimoto (1988). It should be noted that our approach "looks at" the even and odd degree anomalies in an equal weight, while the weight lies entirely or heavily on the even-degree anomalies for the free oscillation or surface wave approach. The  $l=2$  anomaly must be a more distinct anomaly than the adjacent odd degree anomalies. The fast anomaly beneath the East Pacific Rise of our model is not reliable at this depth (see Fig. 8b), but the surface wave study of Suetsugu and Nakanishi (1987) also indicates such a fast anomaly. This fast anomaly extends to the lower mantle at depths up to 1435 km, which seems to be reliable according to the large scale resolution map at depths  $> 800$  km.

At the depth range 1203–1435 km (Fig. 11c), three large scale slow anomalies exist beneath the middle of the Pacific Plate, the Indo–Australia Plate and the African Plate. The lower mantle beneath the East Pacific Rise is generally fast. However, the region beneath other mid oceanic ridges have no systematic anomalies. The global

pattern in general coincides to the model given by Morelli and Dziewonski (1986) shown in Dziewonski and Woodhouse (1987), who retrieved a harmonic expansion of lateral heterogeneity of P-wave velocities from the ISC data set. We notice here that there are fast anomalies along the extension of the great subduction slabs such as the Tonga–Kermadec, Java, South America, and Japan–Izu slabs. One beneath the South America slab is also pointed out by Hager and Clayton (1989). These anomalies are observed from the transition zone down to this depth, suggesting the existence of the lower mantle slab penetrations (Creager and Jordan, 1984). The small scale resolution beneath these areas is generally good (see Fig. 8c).

Figure 11d shows the layer 2566–2900 km just above the CMB. Three major slow anomalies are distinctive, beneath the East Pacific Rise, the subduction region in southwest Pacific, and from the southern Indian Ocean to southwestern Africa. The arctic area including the Bering Sea and Alaska is also slow. The large fast anomaly belt exists from northern Europe to the United States, through the northern Pacific. Antarctica is also fast. These features are in good agreement with the model of Dziewonski (1984) and the model of Morelli and Dziewonski (1986) presented in Giardini et al. (1987).

TABLE 3

Slowness perturbations in each layer. Those for the first iteration are given in parentheses

Layer index	Depths range (km)	Slowness perturbation (%)		
		Min.	Max.	r.m.s.
1	0– 29	– 3.781 (– 1.704)	2.979 (1.219)	0.581 (0.278)
2	29– 78	– 4.138 (– 1.955)	3.540 (1.326)	0.680 (0.308)
3	78– 148	– 4.256 (– 2.019)	3.289 (1.920)	0.705 (0.325)
4	148– 238	– 2.461 (– 1.352)	2.065 (1.216)	0.555 (0.300)
5	238– 348	– 1.628 (– 1.012)	1.783 (1.003)	0.418 (0.248)
6	348– 478	– 1.609 (– 0.775)	1.805 (0.944)	0.365 (0.196)
7	478– 629	– 1.690 (– 0.729)	1.713 (0.914)	0.380 (0.214)
8	629– 800	– 1.452 (– 0.803)	1.189 (0.694)	0.321 (0.181)
9	800– 991	– 1.736 (– 0.957)	1.159 (0.668)	0.313 (0.183)
10	991–1203	– 1.521 (– 0.911)	1.159 (0.851)	0.265 (0.167)
11	1203–1435	– 0.883 (– 0.587)	1.231 (0.686)	0.243 (0.147)
12	1435–1688	– 0.971 (– 0.593)	1.024 (0.568)	0.225 (0.140)
13	1688–1960	– 0.885 (– 0.677)	0.841 (0.425)	0.206 (0.133)
14	1960–2253	– 0.717 (– 0.400)	0.787 (0.493)	0.202 (0.131)
15	2253–2566	– 0.686 (– 0.509)	0.957 (0.554)	0.249 (0.163)
16	2566–2900	– 0.993 (– 0.633)	0.917 (0.594)	0.283 (0.191)

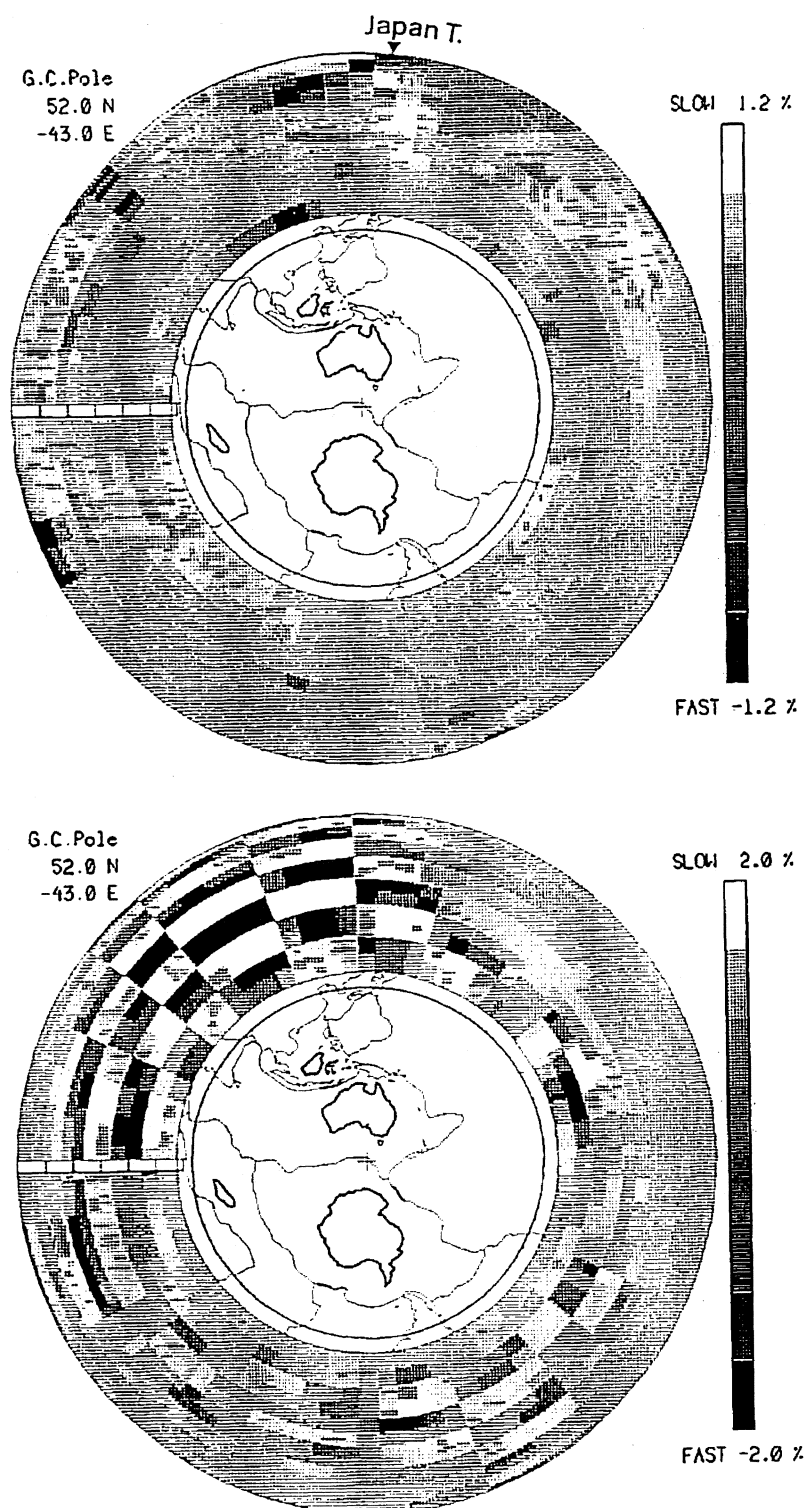


Fig. 12. Vertical cross section of the slowness perturbation (upper) and the checkerboard resolution (lower) along the great circle through Japan and south Africa drawn on the map in the center. Depth scale with 500 km interval is given at the left. Pole position of the great circle is given at the top left.

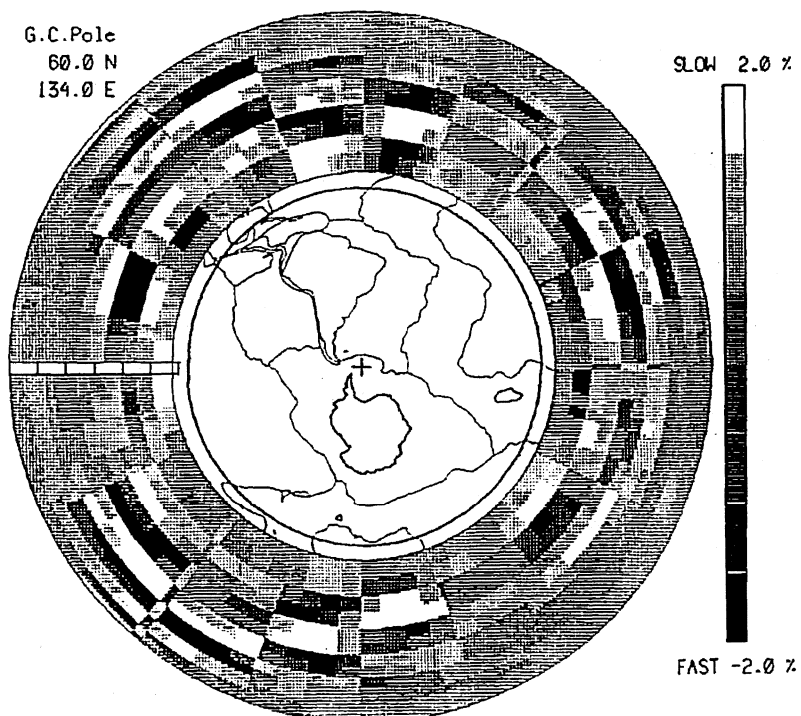
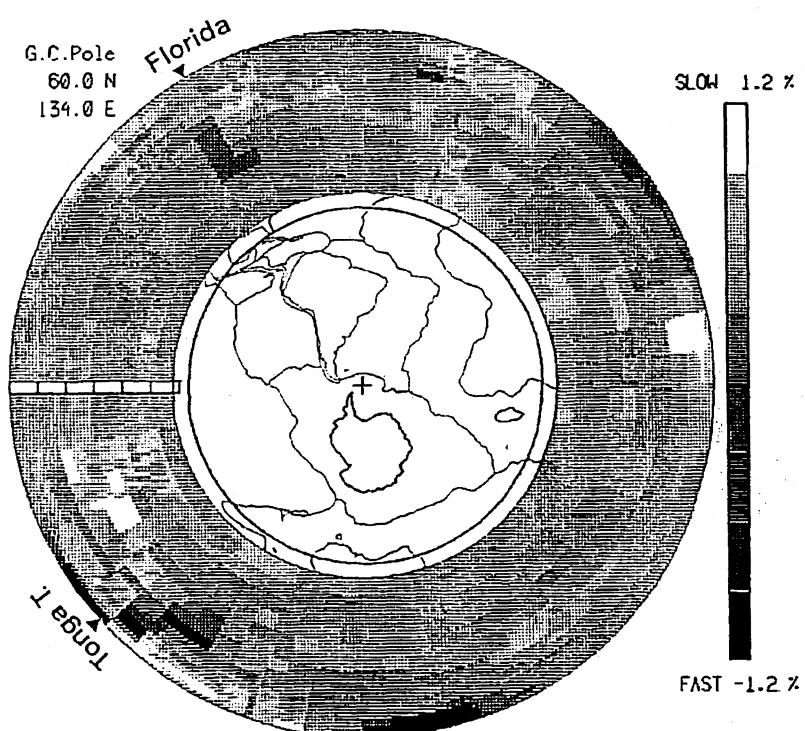


Fig. 13. Vertical cross section along the Gulf of Mexico and the Tonga-New Hebrides trench. See Fig. 12 for other explanations.



The minimum, the maximum and the r.m.s. values of the perturbations in each layer are listed in Table 3. The r.m.s. perturbation is large in the upper mantle, and it decreases with depth down to 1960–2253 km, and again increases towards the CMB. This trend agrees well with the result of Dziewonski (1984). The largest r.m.s. perturbation of 0.705% is in the upper mantle layer 78–148 km shown in Fig. 11a. The smallest value 0.202% is in the layer of 1960–2253 km in the lower mantle. The maximum absolute values of the perturbations in the depth ranges 629–800 km, 1960–2253 km and 2566–2900 km are 1.45%, 0.79% and 0.99%, respectively. In the model of Dziewonski (1984), the minimum r.m.s. perturbation appearing at the depth of 2000 km is three times as small as those at 670 km and CMB. However, our model does not show such a large difference. The large amplitudes in the model of Dziewonski (1984) at 670 km and CMB should be due partly to a boundary effect of the expansion with the Legendre polynomials.

Figures 12 and 13 show the vertical cross-sections of the mantle. The top figures (Fig. 12a and 13a) display the slowness perturbations from the surface to the CMB under the great circle, which is indicated by the thick circle in the map given in the portion of the core. Perturbations greater than  $\pm 1.2\%$  are clipped. The bottom figures (Fig. 12b and 13b) are the checkerboard resolution maps for perturbations of  $\pm 2\%$ ; a black or a white area involves  $4 \times 4 \times 2$  blocks, which is the middle size of the pattern between the top and the bottom of Fig. 8.

In the top part of Fig. 12a, a dipping fast anomaly of up to 1.5% in the upper mantle extending downward from the Japan trench is clearly seen. It extends to a depth of 1000 km beneath the east coast of China. The anomaly is about 2000 km long from the Japan trench; the dip angle is  $25\text{--}30^\circ$  from the surface, in good agreement with the dip angle of the Wadati–Benioff zone. The resolution in this part appears to be quite good according to Fig. 12b. The fast anomaly, thus, strongly suggests the subducting Pacific Plate penetrating into the lower mantle, which is also pointed out in the P-time tomography beneath the Japan Islands by Kamiya et al. (1988). There is a

large, slow anomaly above this dipping fast anomaly, a feature commonly seen in the backarc basins of the Western Pacific. The shallow fast anomaly near the Japan trench is underlain by an exceptional slow anomaly.

Figure 13 shows another section crossing the Tonga trench, Gulf of Mexico, and Africa. Most remarkable is the structure beneath the Tonga trench. Here also exists the dipping fast anomaly of up to 2% extending downward from the Tonga trench. The dip angle is consistent with that of the Wadati–Benioff zone. It reaches a depth of about 1000 km or more beneath the New Hebrides. Also remarkable is an extensive, slow anomaly above the dipping fast anomaly. There is a little slow anomaly overlain by the fast anomaly at the trench, a feature also seen at the Japan trench. Further oceanward there is an exceptional slow anomaly deep in the mantle at depths 1000–2000 km beneath the cluster of oceanic islands such as Society Islands and the Tubuai Islands. Another remarkable structure in this section lies beneath Florida, where a vertically extending fast anomaly of 0.5–1.0% exists at depths from 629 km to 2253 km. This anomaly just coincides to the S-wave velocity anomaly found by Grand (1987). His result shows the fast bodies of about 1% that extend from the depth of 700 km to the deeper limit in his analysis, 1700 km. We support his result and have found that the anomaly extends deeper. A large, slow anomaly extending down to 800 km beneath the Great African Rift Valley is also seen in this section. Although the resolving power is weak around the region, the depth limit is very clear.

### 3.5. Variance mapping

In addition to the checkerboard resolution, the reliability of the solution is tested by calculating approximate variances. Ten independent sets of gaussian noise input have been inverted and the variance of the solution is estimated according to eqn. (33). The standard deviation of the input noise is set to be the same as in the real data inversion, i.e., 0.1, 0.3 and 1.0 s for  $iP$ ,  $P$ , and  $eP$  readings, respectively. The standard deviation of the resultant final residuals should be close to that

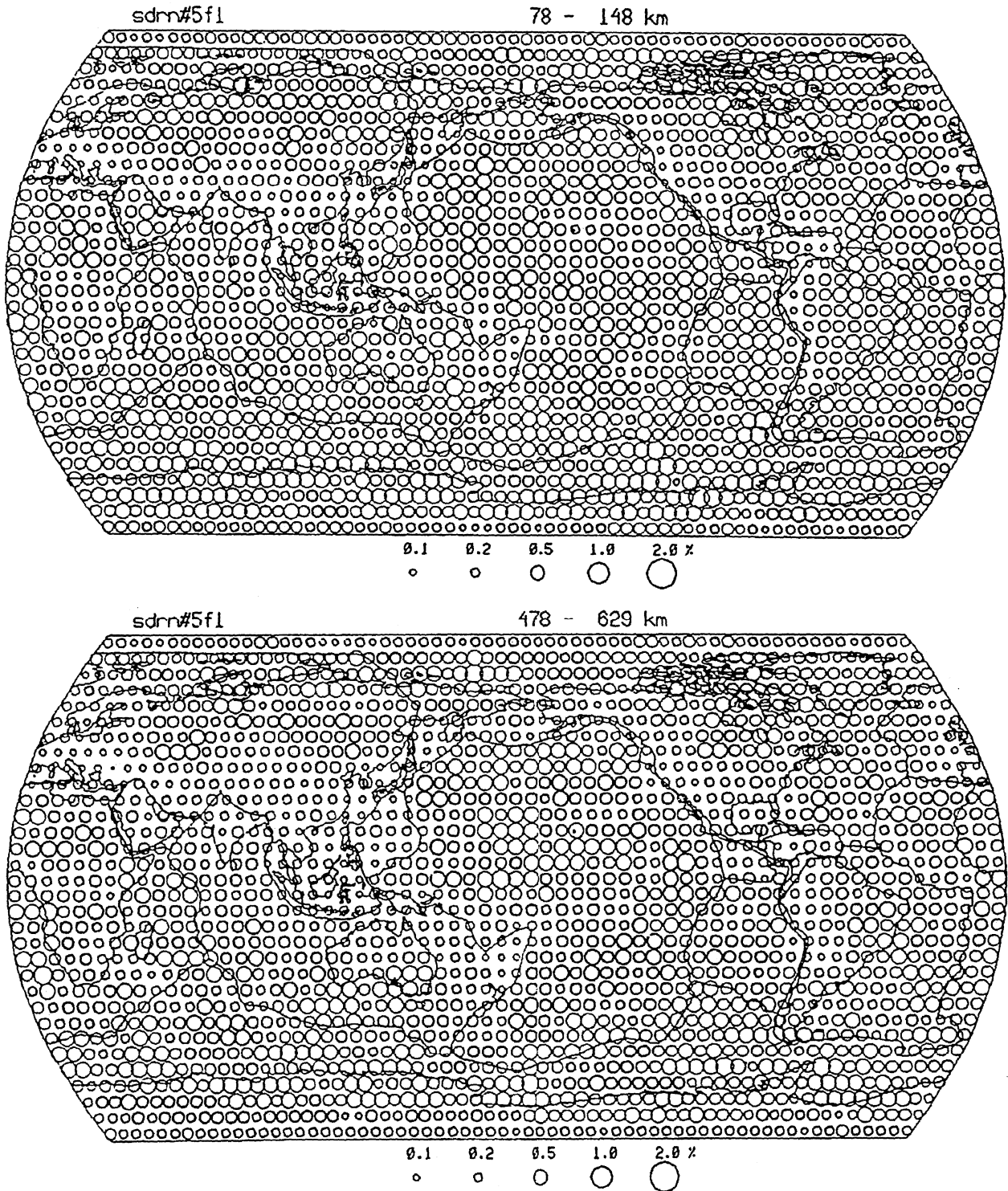


Fig. 14. Variance maps for the data set No. 9 given from gaussian noise input. The corresponding resolution maps and slowness anomaly maps are shown in Figs. 8 and 11, respectively. Size of the circles indicates the standard deviation of perturbations.

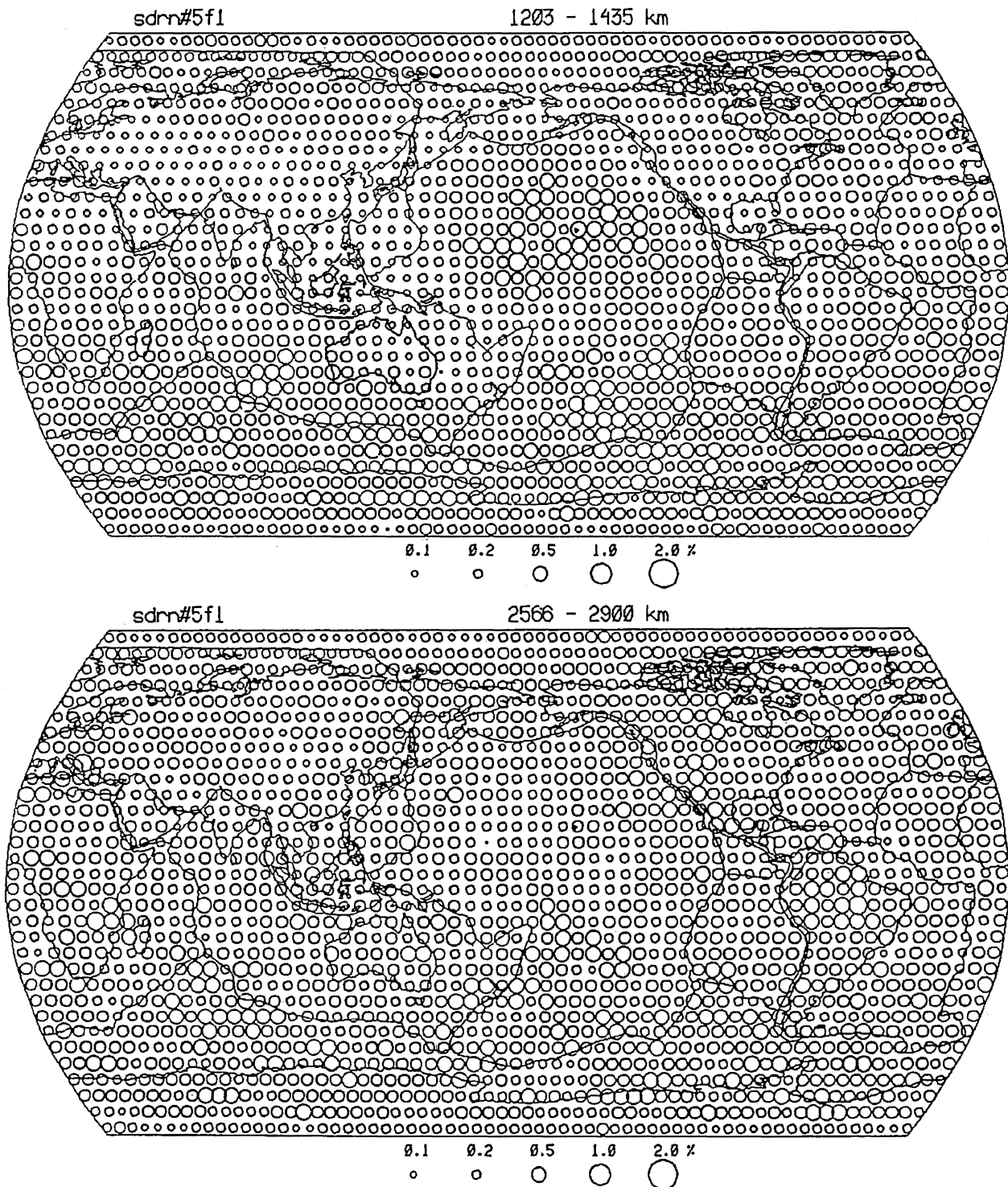


Fig. 14 (continued).

of the reading errors, if our model parameterization is complete, to represent the true Earth. However, the final residuals plotted in Fig. 10b have larger standard deviation of 1.25 s for P readings, which is  $4.2 (= 1.25/0.3)$  times as large as the assumed value. The discrepancy is due probably to both the incomplete parameterization and the wrong assumption on the reading errors. Taking these two possibilities into account, the variance given from the gaussian noise input is rescaled by a factor of 4.2, so that it is of the same magnitude as the variance of the final residuals.

To discuss the variance maps, some comments are deserved. Our final solution has been obtained by stacking the ten solutions. The variance maps shown here are, on the other hand, based on a solution for one data set. Therefore, the plotted values of the standard deviation on the maps must be scaled by a factor of about  $1/\sqrt{10}$ , i.e., the true variance should be much smaller. Figure 14 shows the approximate standard deviations of the solution calculated for the data set of No. 9 at the

same depth ranges as those in Figs. 8 and 11. The gross pattern is the same as that of the resolution maps, i.e., areas of small standard deviation correspond to those resolved well. The solution is accurate in the upper mantle beneath the circum Pacific, South Asia and Europe, and inaccurate beneath the oceans. In the regions of interest, i.e., the areas resolved well, the standard deviations given on the maps are about 0.2–0.5%. Thus, those of the final solution can be estimated to about 0.05–0.2%, which is small enough as compared with the perturbations in the final solution shown in Fig. 11.

Another way of mapping the variance distribution is to calculate the variances of the ten solutions given from the ten real data sets directly; an example is shown in Fig. 15. This provides, in a sense, a more realistic estimate because it is given not from artificial noise but from the real noise on the data. However, it is unfortunate that the introduction of the first order smoothness causes unrealistically small variances at the portion of

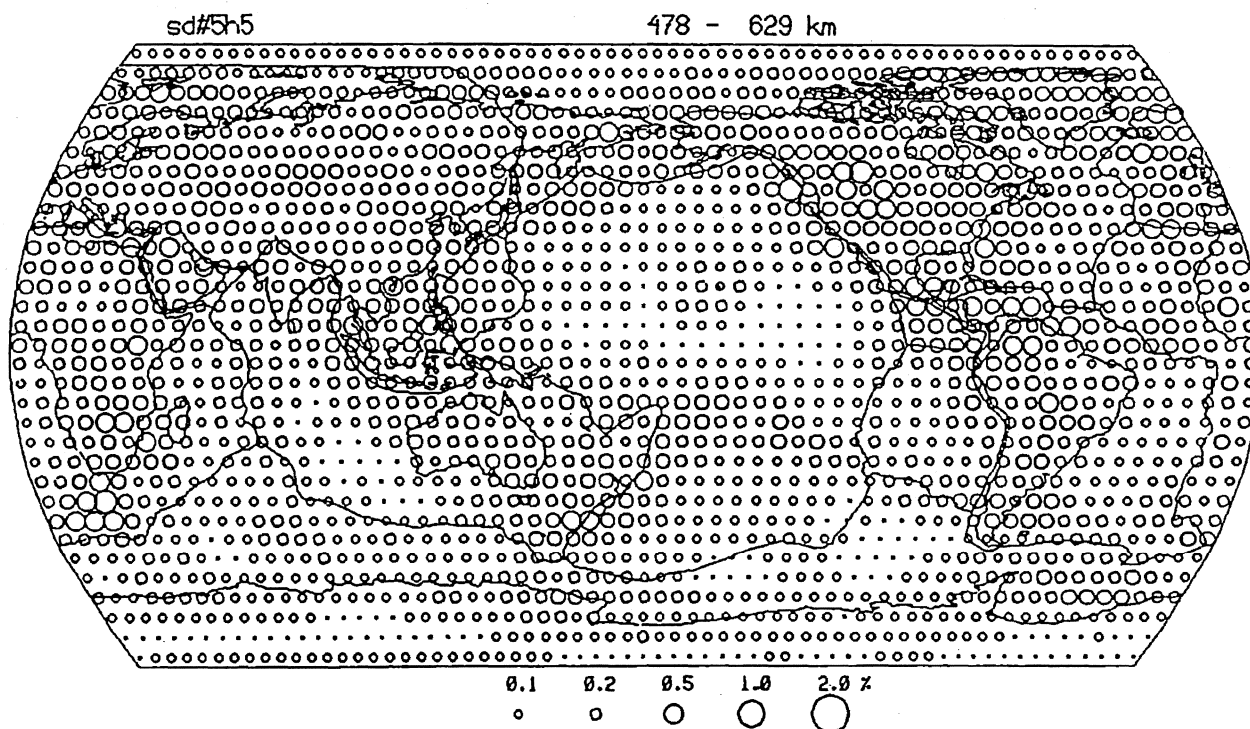


Fig. 15. Variance map given from the ten solutions obtained from the ten data sets in Table 1. Size of the circles indicates the standard deviation of perturbations.

sparse data, such as the upper mantle beneath the oceans. Therefore, the variance cannot be compared from region to region in this map. Only the variance at the portion of abundant data is significant, and the fraction of about  $1/\sqrt{10}$  of the standard deviation gives that of the estimated slowness perturbation. The latter corresponds to the standard deviation estimated from the ten data sets with gaussian noise input. In fact, we see a fair agreement in those values between Figs. 14 and 15 for the areas well resolved.

#### 4. Discussion and conclusion

The global P-time tomographic inversions comparable to this study are those given by Dziewonski (1984) and Hager and Clayton (1989). The largest difference between these and ours is our reconstruction of both the upper and the lower mantle using both shallow and deep events and both near and distant phases. We for the first time have revealed the upper few hundred kilometers of the mantle using P-waves on a global scale. As pointed out by the previous authors, the use of short distance phases is somewhat dubious because of possible phase misidentifications. The result of our inversion might be biased in this sense. On the other hand, their method, in which the upper mantle anomalies are dealt with as station and source anomalies, is also dubious to the same extent. For now it is difficult to estimate either effect quantitatively. Consistency of our result with those from independent data sets for the upper mantle (e.g., Montagner, 1986; Grand, 1987; Suetsugu and Nakanishi, 1987) implies the appropriateness of our approach.

Dziewonski (1984) parameterized the lower mantle with the low order harmonics. To represent such a finer scale heterogeneity as in our study, however, the harmonic expansion would be less adequate than the use of local basis, as far as we invert the body wave delay times. A problem of Dziewonski (1984) is the absence of mapping reliability, though the variances of the harmonic expansion coefficients are tabulated and discussed qualitatively. The problems of Hager and Clayton (1989) may also be pointed out. (1) No iteration

of relocation and backprojection has been done. (2) There is no variance map. (3) The resolution analysis is insufficient. Hager and Clayton (1989) adopted the horizontal block configuration of nearly equal area division which would be better than ours, because our model caused an extremely slow convergence at polar regions in the backprojection. The difference of backprojection technique as well as the number of iterations probably does not affect the gross property of the solution. The largest problem in the studies of Dziewonski (1984) and Hager and Clayton (1989) is the lack of intensive analysis of the reliability of the solutions, which is of primary importance in any field of inverse problems.

Although all three studies thus involve some difficulties, our concern is on the final results. The three results of inversion for the lower mantle are similar in the lowest orders. However, there are also many discrepancies. As discussed by Hager and Clayton (1989), it is not clear at present why the discrepancies arise and which tell the truth. We hope that our result is the best.

We point out here two problems on the vertical structure. First, some systematic trend in the plot of residuals against distance has still remained for each event in the final result, though it is hidden in the superposed plots given in Fig. 10. Second, the vertical roughness parameter  $\sigma_v$  has not been uniquely constrained by the cross validation. A smaller vertical smoothness than one we adopted provides neither better nor worse results. This may imply that the vertical heterogeneity of the mantle relative to our smoothed 1-D model varies more rapidly than allowed by our parameterization, suggesting the potential power of the ISC P time data set to resolve a finer structure of the mantle. There are a number of events in the subduction zones not used in this study. Since the largest anomalies are expected around the upper few hundred kilometers beneath the subduction zones, these areas have to be reconstructed with more finely sampled basis. In terms of computation, the use of finer blocks and larger number of data is not impractical on recent large supercomputers. Although the speed of convergence in the backprojection should be slow for a finely sampled model, there are some techniques to overcome it, such as

successive bisection algorithm adopted in Inoue (1986) or the use of an irregularly sized division with dense grids at the subduction zones.

We presented in this paper the method of whole mantle travel time inversion which we have developed, and the partial result of its application to the vast amount of ISC data. The result for the lower mantle is similar to the previous studies in the longest wavelengths. We, however, resolved the finer structure of both the upper and the lower mantle which should provide ample information for understanding the Earth's global tectonics, such as slab penetration into the lower mantle. The resolution and variance maps obtained will be useful for the interpretation of the result. Full description of the result and its interpretation will be given elsewhere.

### Acknowledgments

We wish to thank Dr. Ken Sudo of International Institute of Seismology and Earthquake Engineering who provided us with facilities of ISC data set. This work was supported by a grant for the cooperative research with the Institute of Statistical Mathematics, the Ministry of Education, Science and Culture of Japan.

### References

- Akaike, H., 1980. Likelihood and Bayes procedure. In: J.M. Bernardo, M.H. De Groot, D.V. Lindley and A.F.M. Smith (Editors), *Bayesian Statistics*, University Press, Valencia, Spain, pp. 143–166.
- Aki, K., Cristoffersson, A. and Husebye, E.S., 1977. Determination of the three-dimensional seismic structure of the lithosphere. *J. Geophys. Res.*, 82: 277–296.
- Aki, K. and Richards, P.G., 1980. *Quantitative Seismology: Theory and Method*. W.H. Freeman, New York, 932 pp.
- Backus, G.E. and Gilbert, F., 1968. The resolving power of gross Earth data. *Geophys. J. R. Astron. Soc.*, 16: 169–205.
- Clayton, R.W. and Comer, R.P., 1983. A tomographic analysis of mantle heterogeneities from body wave travel time data (abstract). *EOS Trans. AGU*, 64: 776.
- Craven, P. and Wahba, G., 1979. Smoothing noisy data with spline functions. *Numer. Math.*, 31: 377–403.
- Creager, K.C. and Jordan, T.A., 1984. Slab penetration into the lower mantle. *J. Geophys. Res.*, 89: 3031–3049.
- Dines, K.A. and Lytle, R.J., 1979. Computerized geophysical tomography. *Proc. IEEE*, 67: 1065–1073.
- Dziewonski, A.M., 1984. Mapping the lower mantle. *J. Geophys. Res.*, 89: 5929–5952.
- Dziewonski, A.M. and Gilbert, F., 1976. The effect of small aspherical perturbations on travel times and a re-examination of the corrections for ellipticity. *Geophys. J.R. Astron. Soc.*, 44: 7–16.
- Dziewonski, A.M. and Anderson, D.L., 1983. Travel times and station corrections for P waves at teleseismic distances. *J. Geophys. Res.*, 88: 3295–3314.
- Dziewonski, A.M. and Woodhouse, J.H., 1987. Global images of the Earth's Interior. *Science*, 236: 37–48.
- Dziewonski, A.M., Hager, B.H. and O'Connell, R.J., 1977. Large-scale heterogeneities in the lower mantle. *J. Geophys. Res.*, 82: 239–255.
- Franklin, J.N., 1970. Well-posed stochastic extensions of ill-posed linear problems. *J. Math. Anal. Appl.*, 31: 682–716.
- Giardini, D., Li, X. and Woodhouse, J.H., 1987. Three-dimensional structure of the Earth from splitting in free-oscillation spectra. *Nature*, 325: 405–411.
- Gilbert, P.F.C., 1972. Iterative methods for three-dimensional reconstructions of an object from projections. *J. Theor. Biol.*, 36: 105–117.
- Grand, S.P., 1987. Tomographic inversion for shear velocity beneath the North American Plate. *J. Geophys. Res.*, 92: 14065–14090.
- Hager, B.H. and Clayton, R.W., 1989. Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid. In: W.R. Peltier (Editor), *Mantle Convection*. Gordon and Breach, New York (in press).
- Herman, G.T., 1980. *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*. Academic Press, New York, 316 pp.
- Herman, G.T., Lent, A. and Rowland, S.W., 1973. ART: Mathematics and applications, a report on the mathematical foundations and on the applicability to real data of the Algebraic Reconstruction Techniques. *J. Theor. Biol.*, 42, 1–32.
- Herrin, E., Tucher, W., Taggart, J., Gordon, D.W. and Lobdell, J.L., 1968. Estimation of surface focus P travel times. *Bull. Seismol. Soc. Am.*, 58: 1273–291.
- Hestenes, M.R. and Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand.*, 49: 409–436.
- Honda, S. and Tanimoto, T., 1987. Regional 3-D heterogeneities by waveform inversion-application to the Atlantic area. *Geophys. J.R. Astron. Soc.*, 91: 737–753.
- Humphreys, E., Clayton, R.W. and Hager, B.H., 1984. A tomographic image of mantle structure beneath southern California. *Geophys. Res. Lett.*, 11: 621–625.
- Humphreys, E. and Clayton, R.W., 1988. Adaptation of back projection tomography to seismic travel time problems. *J. Geophys. Res.*, 93: 1073–1085.
- Inoue, H., 1986. A least squares smooth fitting for irregularly spaced data: Finite-element approach using the cubic B-spline basis. *Geophysics*, 51: 2051–2066.
- Ivansson, S., 1986. Seismic borehole tomography—theory and computational methods. *Proc. IEEE*, 74: 328–338.

- Jackson, D.D., 1979. The use of a priori data to resolve nonuniqueness in linear inversion. *Geophys. J. R. Astron. Soc.*, 57: 137–157.
- Kammerer, W.J. and Nashed, M.Z., 1972. On the convergence of the conjugate gradient method for singular linear operator equations. *SIAM (Soc. Ind. Appl. Math.) J. Num. Anal.*, 9: 165–181.
- Kamiya, S., Miyatake, T. and Hirahara, K., 1988. How deep can we see the high velocity anomalies beneath the Japan Islands? *Geophys. Res. Lett.*, 15: 328–331.
- McMechan, G.A., 1983. Seismic tomography in boreholes. *Geophys. J. R. Astron. Soc.*, 74: 601–612.
- Masters, G., Jordan, T.H., Silver, P.G. and Gilbert, F., 1982. Aspherical Earth structure from fundamental spheroidal-mode data. *Nature*, 298: 609–613.
- Montagnar, J.P., 1986. 3-dimensional structure of the Indian Ocean inferred from long period surface waves. *Geophys. Res. Lett.*, 13: 315–318.
- Morelli, A. and Dziewonski, A.M., 1986. 3-D structure of the Earth's core inferred from travel time residuals. *EOS*, 67: 311.
- Morelli, A. and Dziewonski, A.M., 1987. The harmonic expansion approach to the retrieval of deep Earth structure. In: G. Nolet (Editor), *Seismic Tomography, with Applications in Global Seismology and Exploration Geophysics*. D. Reidel Publishing Co., Dordrecht, Holland, pp. 251–274.
- Nakanishi, I. and Anderson, D.L., 1984. Aspherical heterogeneity of the mantle from phase velocities of mantle waves. *Nature*, 307: 117–121.
- Nakanishi, I., 1985. Three-dimensional structure beneath the Hokkaido-Tohoku region as derived from a tomographic inversion of P-arrival times. *J. Phys. Earth*, 33: 241–256.
- Nataf, H.C., Nakanishi, I. and Anderson, D.L., 1986. Measurement of mantle wave velocities and inversion for lateral heterogeneities and anisotropy 3.inversion. *J. Geophys. Res.*, 91: 7261–7307.
- Neumann-Denzau, G. and Behrens, J., 1984. Inversion of seismic data using tomographical reconstruction techniques for investigations of laterally inhomogeneous media. *Geophys. J. R. Astron. Soc.*, 79: 305–315.
- Nolet, G., 1985. Solving or resolving inadequate and noisy tomographic systems. *J. Comput. Phys.*, 61: 463–482.
- Paige, C.C. and Saunders, M.A., 1982. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Softw.*, 8: 43–71.
- Scales, J.A., 1987. Tomographic inversion via the conjugate gradient method. *Geophysics*, 52: 179–185.
- Sluis, A. van der and Vorst, H.A. van der, 1987. Numerical solution of large, sparse linear algebraic systems arising from tomographic problems. In: G. Nolet (Editor), *Seismic Tomography, with Applications in Global Seismology and Exploration Geophysics*. D. Reidel Publishing Co., Dordrecht, Holland, pp. 49–84.
- Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions (with discussion). *J. R. Stat. Soc., B*, 36: 111–147.
- Suetsugu, D. and Nakanishi, I., 1985. Tomographic inversion and resolution for Rayleigh wave phase velocities in the Pacific Ocean. *J. Phys. Earth*, 33: 345–368.
- Suetsugu, D. and Nakanishi, I., 1987. Three-dimensional velocity map of the upper mantle beneath the Pacific Ocean as determined from Rayleigh wave dispersion. *Phys. Earth Planet. Inter.*, 47: 205–229.
- Tanabe, K., 1971. Projection method for solving a singular system of linear equations and its applications. *Numer. Math.*, 17: 203–241.
- Tanabe, K., 1974. Characterization of linear stationary iterative processes for solving a singular system of linear equations. *Numer. Math.*, 22: 349–359.
- Tanabe, K., 1977. Conjugate-gradient method for computing the Moore–Penrose inverse and rank of a matrix. *J. Optim. Theory*, 22: 1–23.
- Tanimoto, T., 1988. The 3-D shear wave structure in the mantle by overtone waveform inversion-II. Inversion of X-waves, R-waves and G-waves. *Geophys. J. R. Astron. Soc.*, 93: 321–334.
- Tarantola, A. and Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys. Space Phys.*, 20: 219–232.
- Woodhouse, J.H. and Dziewonski, A.M., 1984. Mapping the upper mantle: Three-Dimensional modeling of earth structure by inversion of seismic waveforms. *J. Geophys. Res.*, 89: 5953–5986.



## **第4回情報論的学習理論ワークショップ** **(Information-Based Induction Sciences: IBIS2001)**

### **予稿集**

**日時:** 2001年7月30日(月)~8月1日(水)

**会場:** 学術情報センター 一ツ橋記念講堂 (東京都千代田区)

**主催:** 電子情報通信学会 情報論的学習理論時限研究専門委員会

**共催:** 情報理論とその応用学会

**協賛:** 電子情報通信学会 (情報理論研究専門委員会、人工知能と知識処理研究専門委員会、コンピューテーション研究専門委員会、ニューロコンピューティング研究専門委員会、パターン認識・メディア理解研究専門委員会)、IEEE IT Japan Chapter、情報処理学会、人工知能学会、日本神経回路学会、日本物理学会、計測自動制御学会、システム制御情報学会、応用統計学会、日本ソフトウェア科学会、データマイニング研究会



# Penalized Logistic Regression Machines: New methods for statistical prediction 2

Kunio Tanabe \*

**Abstract:** Support Vector Machines(Vapnik,1979,'95,'98) have been recognized as powerful method for learning certain structures from data for prediction. Their success is due to intrinsic combination of the quadratic programming models with *Kernel Method*. The machine, however, does not seem to accomodate very well the cases where the mechanism of generating data is largely of stochastic nature. By employing the penalized logistic regression model, we make a statistical attempt to construct multiclass discrimination machines which can handle much noisier stochastic data to be competitive with SVM in such an environment. It is shown that *by penalizing the likelihood in a specific way, we can intrinsically combine the logistic regression model with the kernel methods*. In particular, *a new class of penalty functions and associated normalized projective kernels are introduced to gain a versatile induction power of our learning machines*. The closed formulas are given for the first and second derivatives of the log penalized logistic regression likelihood, whose Hessian matrix is shown to be positive definite and uniformly bounded. *Dual classes of globally convergent learning machines(algorithms) are given for obtaining the optimal parameters for both probabilistic and deterministic prediction*. Analysis of the rate of convergence is given for each class of machines. The type-II(or marginal) likelihood and Generalized Information Criteria are also given in closed form for determining the optimal value of hyperparameters in the model so that the machines have a *due induction capacity to the size and the quality of an available training data set*.

## 1 Multiclass Discrimination Problem

Let us consider the problem of multiclass discrimination given a finite number of training data set  $\{(x_i, c_i)\}_{i=1, \dots, N}$ , where  $x_i$  is a column vector of size  $n$  whose elements may be both continuous and discrete numbers and  $c_i$  takes a value in the finite set  $\{1, 2, \dots, K\}$  of classes. We are concerned with the construction of

- i) a conditional multinomial distribution  $\mathcal{M}(p^*(x))$  of  $c$  given  $x \in R^n$ , where  
 $p^*(x)$  is a predictive probability vector whose  $k$ -th element  $p_k^*(x)$  indicates  
the probability of  $c$  taking the value  $k$ ,  
and also
- ii) prediction function  $c = d^*(x) \in \{1, 2, \dots, K\}$

which can be used respectively as stochastic and deterministic prediction of  $c$  given  $x$ .

There have been extensive literatures for solving the problem. They include, for example, Mangasarian (1965, 1968, 1979, 1994), Vapnik(1979, 1995, 1998), Hand(1982, 1997), McLachlan(1992), Ripley(1994, 19

96), Hastie, Tibshirani and Buja(1994), Bishop(1995), Cherkassky and Mulier(1998), Scholkopf, Burges and Smola, A.J. (1999) to name only a few. The purpose of this paper is to give another look at the problem in order to obtain a new class of computational algorithms for predictors and also to introduce criteria for determining a statistically due induction capacity of the predictors to given training data.

Throughout the paper, we will use the following notations: Let  $e_k \equiv (0, \dots, 0, 1, 0, \dots, 0)^t$  be the  $k$ -th unit column vector of size  $K$ , and let the  $K \times N$  constant matrix  $Y$  be defined by

$$Y \equiv [y_1; y_2; \dots; y_N] \equiv [e_{c_1}; e_{c_2}; \dots; e_{c_N}] \quad (1)$$

whose  $j$ -th column vector  $y_j \equiv e_{c_j}$  indicates which class the data  $c_j$  belongs to.

## 2 Normalized Projective Kernels

Following the idea of the Support Vector Machines (SVMs), we introduce a map

$$\phi(x, \lambda) \equiv \begin{pmatrix} \phi_1(x, \lambda) \\ \phi_2(x, \lambda) \\ \vdots \\ \phi_m(x, \lambda) \end{pmatrix} \quad (2)$$

\*The Institute of Statistical Mathematics, 4-6-7 Minamiazabu, Minatoku, Tokyo 106-8569, Japan Tel 03-5421-8754, Fax 03-3446-1695, e-mail tanabe@ism.ac.jp

from  $R^n$  into  $R^m$ , where  $\lambda$  is a hyperparameter vector. For  $\phi(x, \lambda)$ , we can choose a set  $\{\phi_m(x, \lambda)\}_{i=1,2,\dots,m}$  of arbitrarily many functions. We will drop the argument  $\lambda$  for notational simplicity.

Let  $\Phi$  be the  $m \times N$  constant matrix defined by

$$\Phi \equiv [\phi(x_1); \phi(x_2); \dots; \phi(x_N)] \quad (3)$$

whose  $j$ -th column vector is  $\phi(x_j)$ . In order to gain a versatile induction power of the resulting method we always include the constant function  $\phi_0(x) \equiv \omega$  as a member of the regressors, and let the associated augmented map  $\bar{\phi}(x)$  and the  $(m+1) \times N$  constant matrix  $\bar{\Phi}$  be defined respectively by

$$\bar{\phi}(x) \equiv \begin{pmatrix} \omega \\ \phi(x) \end{pmatrix}, \text{ and } \bar{\Phi} \equiv \begin{bmatrix} \omega 1_N^t \\ \Phi \end{bmatrix} \quad (4)$$

where  $\omega$  is a fixed nonnegative constant to be considered as a hyperparameter.

In order to prepare for the arguments in the subsequent sections, we need to introduce a *Normalized Projective Kernel* function  $\mathcal{K}_\omega(x, y)$  of two arguments  $(x, y) \in R^n \times R^n$ , defined by the bilinear form

$$\begin{aligned} \mathcal{K}_\omega(x, y) &\equiv \bar{\phi}^t(x) \bar{\Sigma}^{-1} \bar{\phi}(y) \\ &\equiv \frac{1}{\sigma^2} (\omega^2 + (\sigma\phi(x) - \omega\mu)^t \Sigma^{-1} (\sigma\phi(y) - \omega\mu)), \\ &\equiv \phi_0^2 + (\phi(x) - \phi_0\mu)^t \Sigma^{-1} (\phi(y) - \phi_0\mu), \end{aligned} \quad (5)$$

of the maps  $\bar{\phi}(x)$  and  $\bar{\phi}(y)$ , where  $\bar{\phi}^t(x)$  is the transpose of the column vector  $\bar{\phi}(x)$  and  $\bar{\Sigma}$  is an  $(m+1) \times (m+1)$  positive definite matrix which is parametrized as

$$\bar{\Sigma} \equiv \begin{pmatrix} \sigma \\ \mu \end{pmatrix} \begin{pmatrix} \sigma \\ \mu \end{pmatrix}^t + \begin{bmatrix} 0 & 0^t \\ 0 & \Sigma \end{bmatrix} \equiv \begin{bmatrix} \sigma^2 & \sigma\mu^t \\ \sigma\mu & \Sigma + \mu\mu^t \end{bmatrix}. \quad (6)$$

by a scalar  $\sigma$ , a column vector  $\mu$  and an  $m \times m$  positive definite matrix  $\Sigma$ , which are considered to be hyperparameters, and  $\phi_0 \equiv \omega/\sigma$ . Note that any positive definite matrix  $\bar{\Sigma}$  can be put in this form, because the middle part of Eq.(6) is an  $m$ -step-premature Cholesky decomposition of  $\bar{\Sigma}$ .

It will be shown in Sections 6 and 7 that with a penalized logistic regression model given in the next section we could work directly with the kernel function without resorting explicitly to the map  $\bar{\phi}(x)$  and the parameter  $\bar{\Sigma}$  themselves. In fact, we only need the  $N \times N$  constant matrix

$$\begin{aligned} \mathcal{K}_\omega^d &\equiv [\mathcal{K}_\omega(x_i, x_j)] \equiv \bar{\Phi}^t \bar{\Sigma}^{-1} \bar{\Phi} \\ &\equiv \phi_0^2 1_N 1_N^t + (\Phi - \phi_0\mu 1_N^t)^t \Sigma^{-1} (\Phi - \phi_0\mu 1_N^t) \end{aligned} \quad (7)$$

and the map  $\kappa_\omega(x)$  from  $R^n$  into  $R^N$  defined by

$$\begin{aligned} \kappa_\omega(x) &\equiv \bar{\Phi}^t \bar{\Sigma}^{-1} \bar{\phi}(x) \\ &\equiv \phi_0^2 1_N + (\Phi - \phi_0\mu 1_N^t)^t \Sigma^{-1} (\phi - \phi_0\mu) \end{aligned} \quad (8)$$

There are many possibilities for choosing the hyperparameter  $\mu$ . A typical choice is  $\mu = N^{-1} \Phi^t 1_N = N^{-1} \sum_{j=1}^N \phi(x_j)$ . We briefly touch on the relationship of the normalized projective kernel to the ordinary (but normalized) kernel  $\mathcal{K}_0(x, y) = \phi^t(x) \Sigma^{-1} \phi(y)$ . In particular, we relate the matrix in Eq.(7) and the function in Eq.(8) to the existing ones. If  $\mu$  is chosen as above, then

$$\mathcal{K}_\omega^d = N\phi_0^2 \Pi + (I_N - \phi_0 \Pi_N) \mathcal{K}_0^d (I_N - \phi_0 \Pi_N),$$

$$\kappa_\omega(x) = \phi_0^2 1_N + (I_N - \phi_0 \Pi_N) (\kappa_0(x) - N^{-1} \phi_0 \mathcal{K}_0^d 1_N)$$

where  $\Pi_N \equiv N^{-1} 1_N 1_N^t$  is the orthogonal projector onto the space spanned by  $1_N$ . These equations can be used as formulas for converting the existing kernels to normalized projective kernels.

Smola, Scholkopf and Muller(1998) showed a correspondence between regularization penalties and kernels in the SVM framework. However, we should not interpret their results as saying that a choice of kernel determines regularization penalty. There still remains a lot of rooms for us to control the way in which kernels and penalties are connected, as was shown in Tanabe(2001). For the origin of the kernel method, see Navikoff(1962), Aizermann Braverman and Rozonoer(1964). See also Hand(1982) and Scholkopf, Burges and Smola(1999) for its recent development. For another way of modifying the kernel function, see Amari and Wu(1999) which employed modification of the form,  $c(x)c(y)\mathcal{K}_0(x, y)$ , where  $c(x)$  is a data-dependent function.

### 3 Penalized Logistic Regression Models

In order to solve the problem mentioned earlier, we introduce the penalized logistic regression model. We assume that the joint probability distribution  $\zeta(x, y)$  of  $(x, y)$  from which the training data is sampled is unknown and that the conditional distribution  $\zeta(y|x)$  of  $y$  given  $x$ , follows the multinomial distribution  $\mathcal{M}(p(x))$  specified by the probability vector

$$p(x) \equiv \begin{pmatrix} p_1(x) \\ p_2(x) \\ \vdots \\ p_K(x) \end{pmatrix} \equiv \hat{p}(f(x)), \quad (9)$$

which is parametrized by the logistic transform  $\hat{p}(f)$ , due to Leonard(1973), of the affine transformation  $f(x)$  of the map  $\phi(x)$ , where

$$\hat{p}(f) \equiv \begin{pmatrix} \hat{p}_1(f) \\ \hat{p}_2(f) \\ \vdots \\ \hat{p}_K(f) \end{pmatrix}, \text{ where } \hat{p}_k(f) \equiv \frac{\exp f_k}{\sum_{i=1}^K \exp f_i}, \quad (10)$$

and

$$f(x) \equiv \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_K(x) \end{pmatrix} \equiv \omega w_0 + W\phi(x) \equiv \bar{W} \bar{\phi}(x), \quad (11)$$

$w_0$  and  $W$  are respectively an  $K \times 1$  parameter vector and an  $K \times m$  parameter matrix to be estimated from the given training data,  $\{(x_i, y_i)\}_{i=1, \dots, N}$  as  $y_i$  given in Eq.(1), and the  $K \times (m+1)$  augmented parameter matrix  $\bar{W} \equiv [w_0; W]$  is introduced for notational convenience.

The parametrization of the conditional probability model by the matrix  $\bar{W}$  is a redundant one and the statistical convention is that one of its row vector (say, the  $K$ -th one) is set to be zero. But we maintain the symmetric parametrization in Eq.(11) for both numerical stability and modeling impartiality which were shown in Tanabe and Sagae(1984, 1992, 1999), Tanabe(1985, 1987), Tanabe, Sagae and Ueda(1988) for a nonparametric Bayesian density estimation.

If the data is completely separable by the map  $f(x)$ , there exists no maximum likelihood estimate of  $\bar{W}$  which maximizes the likelihood function

$$L(\bar{W}) \equiv \prod_{j=1}^N p_{c_j}(x_j) \equiv \prod_{j=1}^N \hat{p}_{c_j}(f(x_j)) \equiv \prod_{j=1}^N \hat{p}_{c_j}(\bar{W} \bar{\phi}(x)) \quad (12)$$

with respect to  $\bar{W}$ . Besides, even in the cases where the maximum likelihood estimate  $\bar{W}^{**}$  exists, overfitting could occur with  $\bar{W}^{**}$ . If this is the case, the learning process of maximizing the likelihood suffers from the phenomenon called 'overlearning'. In order to avoid it and obtain a *due induction(generalization) capacity to the size and the quality of an available training data set*, we introduce a penalty function

$$P_{\text{induct}}(\bar{W}) \equiv \exp\left(-\frac{1}{2} \text{trace } \Gamma \bar{W} \Sigma \bar{W}^t\right) \equiv \exp\left(-\frac{1}{2} \|\Gamma^{\frac{1}{2}} \bar{W} \Sigma^{\frac{1}{2}}\|_F^2\right) \quad (13)$$

$$\equiv \exp\left(-\frac{1}{2} (\|\sigma w_0 + W\mu\|_F^2 + \text{trace } \Gamma W \Sigma W^t)\right), \quad (14)$$

where  $\Gamma$  is an  $K \times K$  positive definite matrix,  $\bar{\Sigma}$ ,  $\sigma$ ,  $\mu$  and  $\Sigma$  are given in Section 2,  $\|\cdot\|_F$  is the Frobenius norm, and  $\|a\|_F^2 \equiv a^t \Gamma a$ . We will call it *induction penalty* in this paper.

We employ the penalized logistic regression(PLR) likelihood

$$PL_\delta(\bar{W}) \equiv L(\bar{W}) P_{\text{induct}}^\delta(\bar{W}) = \prod_{j=1}^N \hat{p}_{c_j}(\bar{W} \bar{\phi}(x)) \exp\left(-\frac{\delta}{2} \text{trace } \Gamma \bar{W} \Sigma \bar{W}^t\right) \quad (15)$$

which is to be maximized for obtaining the optimal parameter value  $\bar{W}^*$  of the model, where  $\delta \in [0, \infty)$  is a balancing parameter introduced for notational convenience, although it could be absorbed in the parameters  $\Gamma$  and/or  $\bar{\Sigma}$ . We have introduced the matrices  $\Gamma$  and  $\Sigma$ , the vectors  $\mu$  and  $\lambda$ , the scalars  $\omega$ ,  $\sigma$  and  $\delta$  as *hyperparameters* for the model so that we can gain a variety of *induction(generalization) capacity* of the resulting machines by controlling them according to the sampling scheme of training data set and also to the prospective situation in which the predictor is to be used. Generally, we determine the values of the hyperparameters by a statistical criteria such as the empirical Bayes method, the maximum Type II likelihood method and *GIC* method as will be discussed in Section 8. *The choice of  $\Gamma$  does not affect the kernel function itself* as will be seen in Sections 6 and 7, but it controls the learning process and hence the induction characteristics of the obtainable predictor. In other words, *the induction penalty does not completely specify the kernel function*. We assume through Section 6 that these hyperparameters are fixed. Some of the hyperparameters could be specified *a priori*. In other words, they could be modeled further to meet the specific situation of the training data and the objective of prediction. See [1] for further discussions on this point.

## 4 Maximum PLR Likelihood

The maximum penalized logistic regression likelihood estimate  $\bar{W}^*$  is given by minimizing the negative-log-penalized-likelihood,

$$pl_\delta(\bar{W}) \equiv -\log PL_\delta(\bar{W}) = -\sum_{j=1}^N \log \hat{p}_{c_j}(f(x_j)) + \frac{\delta}{2} \text{trace } \Gamma \bar{W} \Sigma \bar{W}^t. \quad (16)$$

A tedious but straightforward calculation gives the first and the second derivatives of the function  $pl_\delta(\bar{W})$ . Bishop(1992) gave a set of formulas to be composed for obtaining these derivatives, but his formulas can be further simplified to the following closed formulas, in which many terms in his formulas have been cancelled out one another in the case of the likelihood of the multinomial distribution. See also Tanabe and Sagae(1984,'92,'99) in which we employed a similar penalized likelihood but with a positive *semidefinite* matrix  $\Gamma$  for nonparametric Bayesian density estimation. **Lemma 1:** The following equalities hold for  $\delta \geq 0$ .

$$\nabla pl_\delta(\bar{W}) \equiv (P(\bar{W}) - Y) \bar{\Phi}^t + \delta \Gamma \bar{W} \Sigma, \quad (17)$$

$$\nabla^2 pl_\delta(\bar{W}) \equiv \sum_{j=1}^N \{(\bar{\phi}(x_j) \bar{\phi}^t(x_j)) \otimes ([p(x_j)] - p(x_j)p^t(x_j))\} + \delta \bar{\Sigma} \otimes \Gamma \quad (18)$$

where  $\nabla \equiv \partial/\partial \bar{W}$  arranged in the same  $K \times (m+1)$

matrix form as  $\bar{W}$  itself,  $Y$  and  $\bar{\Phi}$  are constant matrices defined respectively in Eqs.(1) and (4),  $P(\bar{W})$  is the  $K \times N$  matrix defined by

$$P(\bar{W}) \equiv [p(x_1); p(x_2); \dots; p(x_N)], \quad (19)$$

whose  $j$ -th column vector is given by

$$p(x_j) \equiv \hat{p}(f(x_j)) \equiv \hat{p}(\bar{W} \bar{\Phi}(x_j)) \quad (20)$$

$[p] \equiv \text{diag}(p)$  is the diagonal matrix formed from the vector  $p$ , and  $\otimes$  is the tensor product.

**Lemma 2:** The following equalities hold for  $\delta \geq 0$ .

$$\begin{aligned} \delta \bar{\Sigma} \otimes \Gamma &< \nabla^2 p_\delta(\bar{W}) \\ &< (\bar{\Phi} \bar{\Phi}^t) \otimes [\forall p] + \delta \bar{\Sigma} \otimes \Gamma \\ &< (\bar{\Phi} \bar{\Phi}^t) \otimes (I_K - K^{-1} 1_K 1_K^t) + \delta \bar{\Sigma} \otimes \Gamma \\ &< (\bar{\Phi} \bar{\Phi}^t) \otimes I_K + \delta \bar{\Sigma} \otimes \Gamma, \end{aligned} \quad (21)$$

where  $A < B$  implies that  $B - A$  is a nonnegative definite matrix,  $I_K$  is the identity matrix of size  $K$ ,  $\Pi_N \equiv N^{-1} 1_N 1_N^t$  is the orthogonal projector onto the space spanned by  $1_N$ , and  $\forall p \equiv \forall p(\bar{W}) \equiv \sum_{j=1}^N p(x_j)$  is the smallest vector such that  $\forall p \geq p(x_j)$  for  $j = 1, 2, \dots, N$  and the inequality is meant elementwise. See [1] for the proof.

**Proposition 3:** The functions,  $p_0(\bar{W})$  and  $p_\delta(\bar{W})$  ( $\delta > 0$ ) are convex and strictly convex functions respectively with respect to the parameter  $\bar{W}$ . The function  $p_\delta(\bar{W})$  has the unique minimum point  $\bar{W}^*$  which satisfies the condition,  $\nabla p_\delta(\bar{W}^*) = O_{K,m+1}$ , where  $O_{K,m+1}$  is the  $K \times (m+1)$  zero matrix. See [1] for the proof.

Now we can give the predictors which will be called *Penalized Logistic Regression Predictors* (PLRPs).

**PLRP:** We adopt  $p^*(x) \equiv \hat{p}(f^*(x)) \equiv \hat{p}(\bar{W}^* \bar{\Phi}(x))$ , as a our predictive probability vector, and  $y^*(x) \equiv \arg \max_k f_k^*(x)$  as our deterministic prediction function, where  $f^*(x) \equiv \bar{W}^* \bar{\Phi}(x)$ .

## 5 Penalized Logistic Regression Machines

Based on Lemma 1 and 2, we can now construct iterative algorithms for computing  $\bar{W}^*$ , which will be called *penalized logistic regression machine* (PLRM). We will give three machines to be chosen by users according to the availability of their computational resources.

**PLRM-1:** Starting with an  $K \times (m+1)$  matrix  $\bar{W}^0$ , generate a sequence of matrices  $\{\bar{W}^i\}_{i=1,2,\dots}$  by the iterative formula for  $i=0,1,2,\dots, \infty$ ,

$$\bar{W}^{i+1} = \bar{W}^i - \alpha_i ((P(\bar{W}^i) - Y) \bar{\Phi}^t + \delta \Gamma \bar{W}^i \bar{\Sigma}), \quad (22)$$

where  $\bar{\Phi}$  is the constant matrix given in Eqs. (3).

**Theorem 4:** The sequence generated by PLRM-1 converges to the unique minimizer  $\bar{W}^*$  of  $p_\delta$  for any choice of initial matrix  $\bar{W}^0$  under certain condition which was specified in Tanabe(2001), for example  $\alpha_i = (\|\bar{\Sigma}\|(\|\mathcal{K}_\omega^d\| + \delta\|\Gamma\|))^{-1}$ . Then its convergence rate is less than  $1 - \delta((\|\mathcal{K}_\omega^d\| + \delta\|\Gamma\|)\|\Gamma^{-1}\|\text{cond } \bar{\Sigma})^{-1}$ , where  $\text{cond } A \equiv \|A\| \|A^{-1}\| (\geq 1)$  is the condition number of a matrix  $A$  and  $\|A\|$  is the spectral norm of  $A$ . See [1] for details and the proof.

**PLRM-1.2:** Starting with an  $K \times (m+1)$  matrix  $\bar{W}^0$ , generate a sequence of matrices  $\{\bar{W}^i\}_{i=1,2,\dots}$  by the iterative formula,

$$\bar{W}^{i+1} = \bar{W}^i - \alpha_i \Delta \bar{W}^i, \quad i = 0, 1, 2, \dots, \infty, \quad (23)$$

where the matrix increment  $\Delta \bar{W}^i$  is the unique solution of the linear matrix equation,

$$\Theta \Delta \bar{W}^i \bar{\Phi} \bar{\Phi}^t + \delta \Gamma \Delta \bar{W}^i \bar{\Sigma} = (P(\bar{W}^i) - Y) \bar{\Phi}^t + \delta \Gamma \bar{W}^i \bar{\Sigma}, \quad (24)$$

where  $\Theta$  is either of the nonnegative definite matrices  $[\forall p]$ ,  $I_K - K^{-1} 1_K 1_K^t$  or  $I_K$ .

We also propose a PLRM, which has a superlinear converge property.

**PLRM-2:** Starting with an  $K \times (m+1)$  matrix  $\bar{W}^0$ , generate a sequence of matrices  $\{\bar{W}^i\}_{i=1,2,\dots}$  by the iterative formula,

$$\bar{W}^{i+1} = \bar{W}^i - \alpha_i \Delta \bar{W}^i, \quad i = 0, 1, 2, \dots, \infty, \quad (25)$$

where  $\Delta \bar{W}^i$  is the unique solution of the linear matrix equation,

$$\begin{aligned} \sum_{j=1}^N ([p(x_j)] - p(x_j)(p(x_j))^t) \Delta \bar{W}^i (\bar{\Phi}(x_j) \bar{\Phi}^t(x_j)) \\ + \delta \Gamma \Delta \bar{W}^i \bar{\Sigma} = (P(\bar{W}^i) - Y) \bar{\Phi}^t + \delta \Gamma \bar{W}^i \bar{\Sigma}. \end{aligned} \quad (26)$$

For the convergence analysis of PLRM-1.2 and PLRM-2 see [1].

## 6 Dual Penalized Logistic Regression Likelihood

The success of SVMs is largely due to intrinsic combination of their quadratic programming model with the kernel methods. We will show that the penalized logistic regression model also yields a certain duality which leads intrinsically to the kernel methods.

Eq.(?) implies that the minimizer  $\bar{W}^*$  of  $p_\delta(\bar{W})$  is of the form

$$\bar{W}^* = V^* \bar{\Phi}^t \bar{\Sigma}^{-1}, \quad (27)$$

where  $V^* \equiv \delta^{-1}\Gamma^{-1}(Y - P(\bar{W}^*))$ . Therefore, by introducing the *dual parameter* matrix  $V$  of size  $K \times N$  in such a way as

$$\bar{W} = V\bar{\Phi}^t\bar{\Sigma}^{-1}, \quad (28)$$

we only have to minimize the negative log penalized logistic regression likelihood

$$p\bar{l}_\delta(V) \equiv pl_\delta(\bar{W}) \equiv pl_\delta(V\bar{\Phi}^t\bar{\Sigma}^{-1}) \quad (29)$$

with respect to  $V$  instead of matrix  $\bar{W}$ . This transformation of the parameters naturally leads to the kernel methods. Substituting Eq.(28) into Eq.(11), we have

$$f(x) \equiv \bar{W} \bar{\phi}(x) = V\bar{\Phi}^t \bar{\Sigma}^{-1} \bar{\phi}(x) = V\kappa_\omega(x). \quad (30)$$

The matrix  $P(\bar{W})$  is unchanged with this transformation, but can also be denoted by  $\tilde{P}(V) \equiv P(\bar{W}) \equiv P(V\bar{\Phi}^t\bar{\Sigma}^{-1})$ , whose  $j$ -th column vector

$$p(x_j) \equiv \hat{p}(f(x_j)) = \hat{p}(V\kappa_\omega(x_j)) \quad (31)$$

can be computed by using the map  $\kappa_\omega(x)$  only. Since the parameter transformation is linear, the function  $p\bar{l}_\delta(V)$  is also a convex function with respect to the *dual parameter*  $V$ , in terms of which it is represented by

$$\begin{aligned} p\bar{l}_\delta(V) &\equiv pl_\delta(V\bar{\Phi}^t\bar{\Sigma}^{-1}) \\ &= -\sum_{j=1}^N \log \hat{p}_{c_j}(V\kappa_\omega(x_j)) + \frac{\delta}{2} \text{trace} \Gamma V \mathcal{K}_\omega^d V^t. \end{aligned} \quad (32)$$

**Remark:** The transformed negative log penalized likelihood  $p\bar{l}_\delta(V)$  involves only the matrix  $\mathcal{K}_\omega^d$ , its column vectors  $\{\kappa_\omega(x_j)\}$  and  $\Gamma$ . Also the predictor probability vector  $\{p(x_j)\}$  involves only the kernel function,  $\kappa_\omega(x)$ . They do not depend explicitly on  $\bar{\Phi}$ ,  $\bar{\phi}(x)$  nor  $\bar{\Sigma}$ .

**Lemma 5:** The derivatives of  $p\bar{l}_\delta$  with respect to the dual parameter  $V$  are given as follows.

$$\nabla p\bar{l}_\delta(V) \equiv (\tilde{P}(V) - Y + \delta \Gamma V) \mathcal{K}_\omega^d, \quad (33)$$

$$\begin{aligned} \nabla^2 p\bar{l}_\delta(V) &\equiv \sum_{j=1}^N \{(\kappa_\omega(x_j) \kappa_\omega^t(x_j)) \\ &\quad \otimes ([p(x_j)] - p(x_j)p^t(x_j))\} + \delta \mathcal{K}_\omega^d \otimes \Gamma \end{aligned} \quad (34)$$

where  $Y$  and  $\mathcal{K}_\omega^d$  are constant matrices given respectively in Eqs.(1) and (7). The second derivatives are uniformly bounded. See [1] for more details and the proof.

**Proposition 6:** There exists at least a minimizer  $V^{**}$  of the negative log penalized likelihood function  $p\bar{l}_\delta(V)$  which satisfies  $\nabla p\bar{l}_\delta(V^{**}) = O_{K,N}$ .

## 7 Dual Penalized Logistic Regression Machines

We can construct an iterative algorithm for computing  $V^{**}$ , which will be called *dual penalized logistic regression machine* (dPLRM).

If the matrix  $\mathcal{K}_\omega^d$  is nonsingular,  $V^{**} = V^*$  is the minimizer which is the unique solution of the matrix equation,

$$D(V) \equiv \tilde{P}(V) - Y + \delta \Gamma V = O_{K,N}. \quad (35)$$

Firstly we give an algorithm for this case.

**dPLRM-0:** Starting with an  $K \times N$  matrix  $V^0$ , generate a sequence of matrices  $\{V^i\}_{i=1,2,\dots}$  by the iterative formula for  $i=0,1,2,\dots, \infty$

$$V^{i+1} = V^i - \alpha_i (P(V^i) - Y + \delta \Gamma V^i). \quad (36)$$

**Theorem 7:** The sequence generated by dPLRM-0 converges to the unique minimizer  $V^*$  of  $p\bar{l}_\delta(V)$  for any choice of initial matrix  $V^0$  under certain condition, which was specified in Tanabe(2001), for example  $\alpha_i = (\|\mathcal{K}_\omega^d\| + \delta\|\Gamma\|)^{-1}$ . Then its convergence rate is less than  $1 - \delta(\|\mathcal{K}_\omega^d\| + \delta\|\Gamma\|)^{-1}$ . See [1] for the proof.

**Remark :** The vector field defined by  $D(V)$  is not a gradient vector field. Since the dual machine dPLRM-0 which employs this vector field is such a simple process as to require only the evaluations of  $\tilde{P}(V) - Y$  and  $\Gamma V$  and no matrix inversion, the present author cannot help speculating if this dual machine is a better approximation to physiological reality of learning process than the existing machines and the other machines introduced in this paper.

If  $\mathcal{K}_\omega^d$  is a singular matrix, then we have the following algorithm, whose convergence is generally slower than dPLRM-0.

**dPLRM-1:** Starting with an  $K \times N$  matrix  $V^0$ , generate a sequence of matrices  $\{V^i\}_{i=1,2,\dots}$  by the iterative formula for  $i=0,1,2,\dots, \infty$ ,

$$V^{i+1} = V^i - \alpha_i (P(V^i) - Y + \delta \Gamma V^i) \mathcal{K}_\omega^d. \quad (37)$$

**Theorem 8:** An accumulating point  $V^\infty$  of the sequence generated by dPLRM-1 with any choice of initial matrix  $V^0$  is a minimizer of  $p\bar{l}_\delta(V)$  under certain condition, which was specified in Tanabe(2001), for example  $\alpha_i = \|\mathcal{K}_\omega^d\|^{-1}(\|\mathcal{K}_\omega^d\| + \delta\|\Gamma\|)^{-1}$ . Then its convergence rate is less than  $1 - \delta(\|\mathcal{K}_\omega^d\| + \delta\|\Gamma\|)^{-1}$ . See [1] for the proof.

We also give a dPLRM which has a better linear converge property.

**dPLRM-1.2:** Starting with an  $K \times N$  matrix  $V^0$ , generate a sequence of matrices  $\{V^i\}_{i=1,2,\dots}$  by the iterative formula,

$$V^{i+1} = V^i - \alpha_i \Delta V^i, \quad i = 0, 1, 2, \dots, \infty, \quad (38)$$

where  $\Delta V^i$  is the solution of the linear matrix equation,

$$\Theta \Delta V \mathcal{K}_\omega^d + \delta \Gamma \Delta V = P(V^i) - Y + \delta \Gamma V^i, \quad (39)$$

where  $\Theta$  is either of the nonnegative definite matrices  $[Vp]$ ,  $I_K - K^{-1}1_K 1_K^t$  or  $I_K$ .

**Remark :** In fact, if  $\mathcal{K}_\omega^d$  is nonsingular, then  $\Delta V^i = (\mathcal{K}_\omega^d \otimes \Theta + \delta I_N \otimes \Gamma)^{-1} (\mathcal{K}_\omega^d \otimes I_K)^{-1} \nabla p_\delta(V^i)$ , whose direct computation is numerically worse than solving Eq.(39).

We also give a dPLRM which has rapid converge property.

**dPLRM-2:** Starting with an  $K \times N$  matrix  $V^0$ , generate a sequence of matrices  $\{V^i\}_{i=1,2,\dots}$  by the iterative formula,

$$V^{i+1} = V^i - \alpha_i \Delta V^i, \quad i = 0, 1, 2, \dots, \infty, \quad (40)$$

where  $\Delta \bar{W}^i$  is the solution of the linear matrix equation,

$$\sum_{j=1}^N ([p(x_j)] - p(x_j)(p(x_j))^t) \Delta V^i (\kappa_\omega(x_j) \kappa_\omega^t(x_j)) + \delta \Gamma \Delta V^i \mathcal{K}_\omega^d = (P(V^i) - Y + \delta \Gamma V^i) \mathcal{K}_\omega^d. \quad (41)$$

For the convergence analysis of dPLRM-1.2 and dPLRM-2 see [1].

**Remark:** We can work out the process of getting the predictor only with the quantities related to the symbols  $V, \bar{P}(V), \mathcal{K}_\omega^d, \kappa_\omega(x)$  and  $\Gamma$  without resorting at all to  $\bar{W}, P(\bar{W}), \bar{\Phi}, \bar{\phi}(x)$  nor  $\bar{\Sigma}$ , which we can also do without for evaluation of the functions,

$$f^*(x) = V^* \kappa_\omega(x), \quad (42)$$

and the predictive probability

$$p^*(x) \equiv \hat{p}(f^*(x)) = \hat{p}(V^* \kappa_\omega(x)), \quad (43)$$

for the prediction given  $x$ . This implies that we could perform the above mentioned learning and prediction process only with the kernel function  $\mathcal{K}(x, y)$  without resort to the original map  $\phi(x)$  itself. It does not even matter whether the kernel function is constructed from such a map. The situation is largely parallel to that of SVM. See Scholkopf et al.(1999). The methods described in this and the previous sections will be called *dual* penalized logistic regression methods and the method described in Sections 4 and 5 will be called *primal* methods. Likewise, the algorithms given in Section 5 will be called *primal PLR machines* as against dual PLR machines.

## 8 Maximum Marginal Likelihood Method for Hyperparameters

For determining the values of the hyperparameters, which produces a proper *induction capacity* given the

training data, we propose the application of the *maximum type II(or Marginal) likelihood method* due to Good(1965) which was rediscovered by Akaike(1980) who proposed the criterion  $ABIC(\delta) \equiv -2 \log L(\delta)$ , where  $L(\delta)$  is the type-II likelihood given by,

$$L(\delta) \equiv \int L(y|\theta) \pi(\theta|\delta) d\theta, \quad (44)$$

$L(y|\theta)$  is the likelihood function,  $\pi(\theta|\delta)$  is a prior,  $\theta$  is a parameter and  $\delta$  is a hyperparameter. Applying this method to the present model with  $\pi(\bar{W}|\delta) \equiv (\frac{\delta}{2\pi})^{K(m+1)/2} \exp(-\frac{\delta}{2} \text{trace} \Gamma \bar{W} \bar{\Sigma} \bar{W}^t)$  we obtain the criteria, ABIC

$$\begin{aligned} \equiv -2pl_\delta(\bar{W}^*) &= \log \det \nabla^2 pl_\delta(\bar{W}^*) \\ &\quad - K(m+1) \log \delta + \text{correction} \end{aligned} \quad (45)$$

$$\begin{aligned} \equiv -2\tilde{pl}_\delta(V^*) &= \log \det \nabla^2 \tilde{pl}_\delta(V^*) - \log \det \mathcal{K}_\omega^d \\ &\quad + \log \det \bar{\Sigma} - K(m+1) \log \delta + \text{correction} \end{aligned} \quad (46)$$

where  $\nabla^2 pl_\delta(\bar{W})$  and  $\nabla^2 \tilde{pl}_\delta(V)$  are given respectively in Eqs.(18) and (34). The correction terms in the formula are computed by the Monte Carlo method which employs importance sampling from the Gaussian approximation with the covariance matrices  $[\nabla^2 pl_\delta(\bar{W}^*)]^{-1}$ . It could be discarded if an approximate value of the criterion is needed. For the detail of the technique for computing the orrection term, see Section 5 of Tanabe and Sagae(1999) which was employed in their works (1984,1985,1987,1988) for Bayesian nonparametric density estimation.

Alternatives criteria for choosing the value of the hyperparameter include *Generalized Information Criterion*(GIC) due to Konishi and Kitagawa(1996) and *Generalized Cross Validation* due to Whaba(1985). See also Konishi(2000) for a general introduction to GIC. The GIC statistic with our model is given, in closed form, by GIC( $\delta$ )

$$\begin{aligned} \equiv -2pl_\delta(\bar{W}^*)|_{\delta=0} + 2\text{trace}[\nabla^2 pl_\delta(\bar{W}^*)]^{-1} \\ (\nabla pl_\delta(\bar{W}^*) \otimes \nabla pl_\delta(\bar{W}^*)|_{\delta=0}), \end{aligned} \quad (47)$$

where  $|_{\delta=0}$  means the value evaluated at  $\delta = 0$  with  $\bar{W}^*$  obtained for positive  $\delta$ ,  $\nabla pl_\delta(\bar{W}^*)$  is given in Eq.(17).

**References** is not included here due to space limitation. But, for the literatures cited in this paper and related works, see the refereces of

[1] Kunio Tanabe, *Penalized Logistic Regression Machines: New methods for statistical prediction 1*, ISM Cooperative Research Report 143, Estimation and Smoothing Methods in Nonparametric Statistical Models, 163-194 (2001)

from which the present paper is summerized.

# Speaker Recognition without Feature Extraction Process

Tomoko MATSUI and Kunio TANABE

The Institute of Statistical Mathematics  
4-6-7 Minami-Azabu, Minato-ku, Tokyo, 106-8569 Japan

E-mail: {tmatsui, tanabe}@ism.ac.jp

**Abstract** By employing the dual Penalized Logistic Regression Machine (dPLRM), this paper explores a speaker identification method which does not require feature extraction process depending on a prior knowledge. The induction machine can discover implicitly speaker characteristics relevant to discrimination only from a set of training data by the mechanism of the kernel regression. Our text-independent speaker identification experiments with training data uttered by 10 male speakers in three different sessions show that the proposed method is competitive with the conventional Gaussian mixture model (GMM)-based method with 26-dimensional Mel-frequency cepstrum (MFCC) feature even though our method handle directly coarse data of 256-dimensional log-power spectrum. It is also shown that our method outperforms the GMM-based method especially as the amount of training data becomes smaller.

**Key words** Speaker recognition, Speaker identification, Kernel regression, dual Penalized Logistic Regression Machine, implicit feature extraction

## 1. Introduction

There have recently been great demands for automatic speaker recognition in such applications as securing a protected access to various information services and indexing speakers in sound archives. As is shown in the series of the NIST reports [1] on annual evaluations of text-independent recognition studies conducted by research laboratories all over the world, the state of the art method is based on modeling individual speakers with GMMs via a set of reduced data of Mel-frequency cepstral coefficients (MFCCs) with a few dozens of the dimension.

While the MFCC data have been known to well capture the psycho-physical characteristics [4] and are widely believed to annihilate unwanted fluctuations in speech of individual speakers [5], there is no reason to assume that some useful information for speaker identification might not be lost in the reduced data. Besides, the dimension of MFCC vectors might have been chosen to accommodate the stable

computation of estimate of the parameters in the GMM. For the filter-bank analysis, Biem et al. reported the method of estimating a better scale for speech recognition based on the discriminative training [6], and Miyajima et al. successfully applied the method to speaker recognition [7]. The Mel-scale, however, might not be needed for speaker identification.

This paper attempts to avoid outright pre-processing of speech data such as the MFC transform by handling directly the coarse data with 256-dimensional log-power spectrum by employing the dual penalized logistic regression machine dPLRM [8-10] which, being applied to the MFCC data, has already been shown to be competitive with the GMM-based method and also with the support vector machine (SVM) [11,12]. While the GMM-based method, which estimates a density function for each speaker independently, requires large amount of training data to learn the characteristics of individual speakers, the dPLRM needs less amount of training data since it can handle nonlinearity

more effectively with kernel functions and do discriminating learning interdependently. As a dual machine of the logistic regression machine, the dPLRM has a versatile expressiveness of hidden structures in the training data and induction power beyond expectation [8-10]. Figure 1 shows the respective speaker identification processes of our dPLRM method and the conventional GMM-based method.

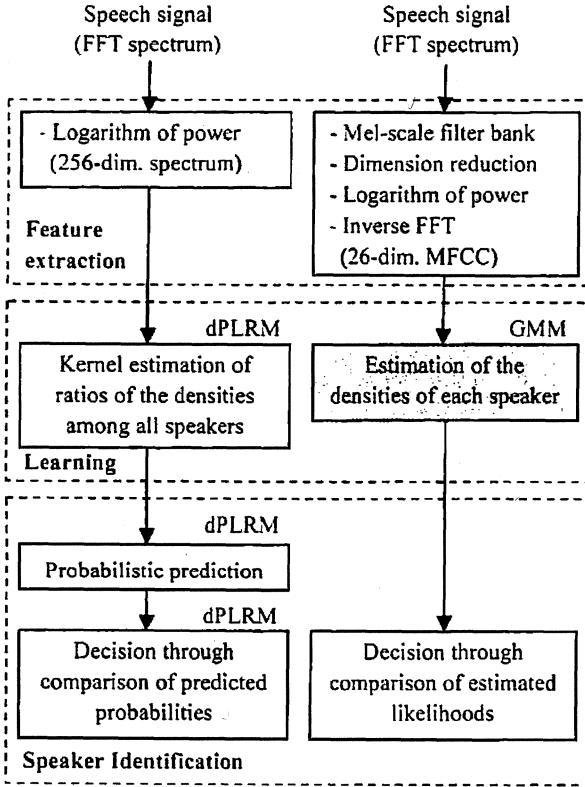


Figure 1. The dPLRM method (the left-hand side) vs. the GMM-based method (the right-hand side).

In Section 2, we briefly sketch the dPLRM method. In Section 3, when it is applied to the training data recorded in different sessions, the dPLRM method is shown to outperform the GMM-based method with MFCCs even though it avoids an elaborate extraction process. We also discuss the case in which the methods are trained on a coarsely sampled data.

## 2. dPLRM for speaker identification

### 2.1 dual Penalized Logistic Regression Machine

Let  $x_j$  is a column vector of size  $n$  and  $c_j$  takes a

value in the finite set  $\{1, 2, \dots, K\}$  of classes. The learning machine dPLRM feeds a finite number of training data  $\{(x_j, c_j)\}_{j=1, \dots, N}$ , and then produces a conditional multinomial distribution  $M(p^*(x))$  of  $c$  given  $x \in \mathbb{R}^n$ , where  $p^*(x)$  is a predictive probability vector whose  $k$ -th element  $p_k^*(x)$  indicates the probability of  $c$  taking the value  $k$ .

For convenience, we code the class data  $c_j$  by  $j$ -th unit column vector  $e_k \equiv (0, \dots, 1, \dots, 0)^T$  of size  $K$  and define an  $K \times N$  constant matrix  $Y$  by

$$Y \equiv [y_1; \dots; y_N] \equiv [e_{c_1}; \dots; e_{c_N}] \quad (1)$$

whose  $j$ -th column vector  $y_j \equiv e_{c_j}$  indicates the class to which the data  $x_j$  is attached.

We introduce a mapping from  $\mathbb{R}^n$  into  $\mathbb{R}^K$ ,

$$F(x) \equiv V k(x) \quad (3)$$

where  $V$  is an  $K \times N$  parameter matrix which is to be estimated by using the training data set  $\{(x_j, c_j)\}_{j=1, \dots, N}$ .

$k(x)$  is a map from  $\mathbb{R}^n$  into  $\mathbb{R}^N$  defined by

$$k(x) \equiv (K(x_1, x), \dots, K(x_N, x))^T, \quad (4)$$

and  $K(x, x')$  is a certain positive definite kernel function. Then we define a multinomial model for probabilistic predictor  $p(x)$  by

$$p(x) \equiv \hat{p}(F(x)) \equiv (\hat{p}_1(F(x)), \dots, \hat{p}_K(F(x)))^T, \quad (5)$$

where  $\hat{p}_k(F(x)) \equiv \frac{\exp(F_k(x))}{\sum_{i=1}^K \exp(F_i(x))}$  is the logistic transform.

Under this model assumption, the negative log-likelihood function  $L(V)$  for  $p(x)$  is given by

$$L(V) \equiv -\sum_{j=1}^N \log(p_{c_j}(x_j)) = -\sum_{j=1}^N \log(\hat{p}_{c_j}(V k(x_j))) \quad (6)$$

which is a convex function. This objective function  $L(V)$  is of discriminative nature, and that if the kernel function is appropriately chosen, the map  $F(x)$  can represent a wide variety of functions so that the resulting predictive probability  $p(x)$  can be expected to be close to the reality. A



predictive vector  $p^*(x)$  could be obtained by putting  $p^*(x) = \hat{p}(V^{**}k(x))$  where  $V^{**}$  is the ML estimate which minimize the function  $L(V)$  with respect to  $V$ .

However, over-learning problems could occur with  $V^{**}$  with the limited number of training data. In order to deal with the problems, the penalty term is introduced and the negative-log-penalized-likelihood

$$PL(V) \equiv L(V) + \frac{\delta}{2} \left\| \Gamma^{\frac{1}{2}} V \bar{K}^{\frac{1}{2}} \right\|_F^2 \quad (7)$$

is minimized to estimate  $V$  where  $\|\cdot\|_F$  is the Frobenius norm. The penalty term is intended to reduce the effective freedom of the variable  $V$ . The matrix  $\Gamma$  is an  $K \times K$  positive definite matrix. A frequent choice of  $\Gamma$  is given by

$$\Gamma = \frac{1}{N} Y Y^T \quad (8)$$

which equilibrates a possible imbalance of classes in the training data. The matrix  $\bar{K}$  is the  $N \times N$  constant matrix, given by

$$\bar{K} = [K(x_i, x_j)]_{i,j=1,\dots,N} \quad (9)$$

The  $\delta$  is a regularization parameter and can be determined by the empirical Bayes method.

Due to the introduction of the specific quadratic penalty in (7), the minimizer  $V^*$  of  $PL(V)$  is a solution of the neat matrix equation.

$$\nabla PL \equiv (P(V) - Y + \delta \nabla V) \bar{K} = O_{K,N}, \quad (10)$$

where  $P(V)$  is an  $K \times N$  matrix whose  $j$ -th column vector is the probability vector  $p(x_j) \equiv \hat{p}(V k(x_j))$ . The matrix  $Y$  is given in (1). The minimizer  $V^*$ , which gives the probabilistic predictor  $p^*(x) \equiv \hat{p}(V^* k(x))$ , is iteratively computed by the following algorithm.

**Algorithm:** Starting with an arbitrary  $K \times N$  matrix  $V^0$ , we generate a sequence  $\{V^i\}$  of matrices by

$$V^{i+1} = V^i - \alpha_i \Delta V^i, \quad i = 0, \dots, \infty \quad (11)$$

where  $\Delta V^i$  is the solution of the linear matrix equation,

$$\sum_{j=1}^N ([p(x_j)] - p(x_j)(p(x_j))^T) \Delta V^i (k(x_j)(k(x_j))^T) + \delta \nabla V^i \bar{K} = (P(V^i) - Y + \delta \nabla V^i) \bar{K}. \quad (12)$$

The detailed algorithm for estimation is shown in [8-10]. Note that we only need to solve an unconstrained optimization of a strictly convex function  $PL(V)$  or equivalently, to solve the simple matrix nonlinear equation (10).

## 2.2 Speaker identification procedure

Figures 2 and 3 show the training and testing procedures respectively.

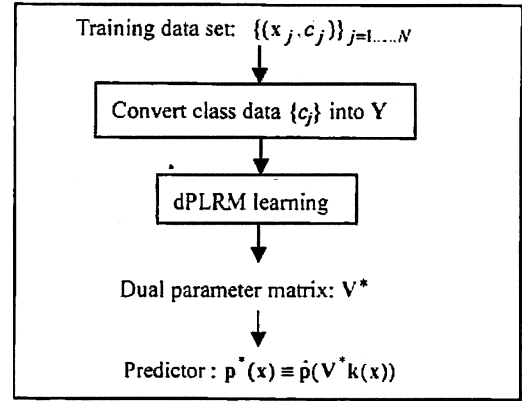


Figure 2. Training procedure.

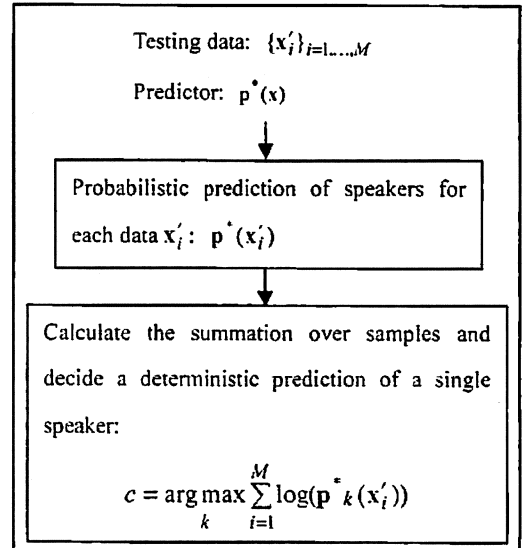


Figure 3. Testing procedure.

### 2.3 Expressiveness of the polynomial kernel function

In the previous section, we have introduced the mapping  $F(x)$  a priori for brevity. In fact, the dPLRM was introduced by Tanabe [8-10] as a dual machine of the penalized logistic regression machine PLRM in which  $F(x)$  is represented by

$$F(x) = W\phi(x) \quad (13)$$

where  $\phi(x) = (\phi_1(x), \dots, \phi_m(x))^t$ , each element of which is a certain nonlinear function of  $x$ . The PLRM [8-10] minimizes the penalized likelihood function

$$PL(W) \equiv L(W) + \frac{\delta}{2} \left\| \Gamma^{\frac{1}{2}} W \Sigma^{\frac{1}{2}} \right\|_F^2, \quad (14)$$

where  $\Sigma$  is a positive definite matrix. It was also shown that dPLRM and PLRM give exactly the same predictor  $p'(x)$  when  $\phi(x)$ ,  $\Sigma$  and  $K(x, x')$  are appropriately chosen and that the former is computationally far less expensive than the latter. For the speaker identification problem we treat with dPLRM in this paper, we make use of the polynomial kernel function

$$\begin{aligned} K(x, x') &= (x^t x' + 1)^s \\ &= \sum_{j=0}^s sCj (x^t x')^j = \sum_{j=0}^s \left[ sCj \left( \sum_{i=1}^n [x]_i [x']_i \right)^j \right] \end{aligned} \quad (15)$$

which is equivalent to the choice of

$$\begin{aligned} \phi(x) &= (1, x_1, \dots, x_{256}, \\ &\quad x_1^2, \dots, x_1 x_2, \dots, \\ &\quad x_1^3, \dots, x_1^2 x_2, \dots, x_1 x_2 x_3, \dots, \\ &\quad x_1^4, \dots, \\ &\quad \dots) \end{aligned} \quad (16)$$

$$\begin{aligned} \Sigma^{-1} &= \text{diag}(1, 1, \dots, 1, \\ &\quad 1, \dots, 2, \dots, \\ &\quad 1, \dots, 3, \dots, 6, \dots, \\ &\quad 1, \dots, \\ &\quad \dots) \end{aligned}$$

in PLRM, where  $sCj$  is the number of combinations of  $s$  taken  $j$  at a time and  $[x]_i$  is the  $i$ -th degree monomial in the elements of  $x \in \mathbb{R}^n$ . If we chose  $s=5$  as is the case with the experiments given in Section 3, the number  $m$  of elements of  $\phi(x)$  is so huge as  $O(10^{10})$ . Therefore it may be easily seen that the expressive power of the map  $F(x)$  is so high that the map could mimic, if necessary, the operations indicated in the feature extraction process of the GMM-based method in Figure 1. Our experiment suggests that without resorting to human judgment such as the Mel-scale filtering, the dPLRM can automatically construct some kind of nonlinear transformation from the training data although it might not be similar to the feature extracting transformation employed in the GMM-based method.

### 3. Experiments

The performances of our method and the GMM-based method are compared through text-independent speaker identification experiments.

#### 3.1 Data description and experimental conditions

The data has been collected for 10 male speakers who utter several sentences (for four seconds per sentence) and words (for one second per word). Although the texts are common for all speakers, the sentences used for testing are different from those for training. The utterances were recorded at the sampling rate of 16 kHz in six sessions from T0 through T5 over 13 months' period. The interval between T0 and T1 is one month and the other intervals are three months. A 256-dimensional log power spectrum vector and a MFCC vector of 26 components, consisting of 12 Mel-frequency cepstral coefficients plus normalized log energy and their first derivatives, is derived once every 10 ms over a 25.6 ms Hamming-windowed speech segment.

We choose two kinds of training data sets, DS1 and DS2. The set DS1 consists of the data for three sentences, each of which is uttered in Session T0, T1 and T2, respectively, and the set DS2 consists of the data for three sentences uttered in the single session T2. The total duration

of the utterances of three sentences is approximately 12 seconds per speaker. For testing purpose, we choose the utterances of the five sentences and the five words from Sessions T3, T4 and T5 and test them individually. For both sentence and word cases, the total case number of the testing is 150 since we have 10 speakers times 5 sentences(or words) and 3 sessions.

The polynomial kernel function (13) is used for the dPLRM. The power is chosen to be  $s=5$  for the log power spectrum data and  $s=9$  for the MFCC data, respectively. The parameters  $\alpha$  and  $\delta$  in dPLRM are experimentally set to be 1.0 in (11) and  $1.3e-3$  in (7), respectively. In order to execute effective computation with 64-bit precision, the data is so scaled that all the elements of feature vectors lie in the interval  $[-0.5, 0.5]$ .

In the GMM-based method, the mixture model of 16 Gaussian distributions with diagonal covariance was chosen as a speaker model among the competing models with 8, 16 and 24 Gaussian distributions. The parameters were initialized using all training speech for all speakers with the HMM toolkit (HTK) [13], and then estimated with the EM algorithm for each speaker.

For testing the methods, the deterministic decision rule given in Figure 3 is adopted with the dPLRM method, although the dPLRM gives generally a probabilistic prediction. On the other hand, the GMM method adopts the decision rule to select the speaker who attained the maximum collective log-likelihood.

### 3.2 Test of DS1-trained methods

Firstly we compare the performances of the methods trained on the set DS1. Table 1 lists the identification rates with the confidence intervals at a confidence level of 90% averaged over the 150 cases.

The dPLRM method with the log power spectrum performed the best for both word and sentence speech. Since the training data contains the information on utterance variations among Sessions T0, T1 and T2, our method attains higher success rates.

Table 1. Speaker identification rates (with the confidence intervals at a confidence level of 90%) using the training data of the MFCCs and log power spectrum extracted from three sentences uttered in Session T0/T1/T2 for each sentence.

Testing data	Method	Identification rates (%)	
		MFCC	Log power spectrum
Word speech	dPLRM		92.7 (89.3, 96.0)
	GMM	89.3 (85.3, 93.3)	84.0 (79.3, 88.7)
Sent. speech	dPLRM		100 (99.3, 100)
	GMM	99.3 (98.7, 100)	99.3 (98.7, 100)

### 3.3 Test of DS2-trained methods

Secondly we test the methods trained on the set DS2. Table 2 lists the identification rates with the same confidence qualification as stated above.

Table 2. Speaker identification rates (with the confidence intervals at a confidence level of 90%) using the training data of the MFCCs and log power spectrum extracted from three sentences uttered in Session T2.

Testing data	Method	Identification rates (%)	
		MFCC	Log power spectrum
Word speech	dPLRM	88.7 (84.7, 92.7)	83.3 (78.7, 88.0)
	GMM	84.7 (80.0, 89.3)	68.0 (62.0, 74.0)
Sent. speech	dPLRM	98.7 (97.3, 100)	97.3 (95.3, 99.3)
	GMM	98.0 (96.0, 99.3)	86.7 (82.7, 91.3)

Both dPLRM and GMM-based methods trained on the MFCC data gave higher identification rates than those trained on the log power spectrum data. We note that the performance with the GMM-based method drops drastically when the training data switches from the MFCCs to the log power spectrum. We found some difficulties with the GMM-based method in the estimation process due to the high dimensionality of the 256-dimensional log power spectrum data.

### 3.4 Test of the methods trained on coarsely sampled data

Table 3 lists the identification rates with the same confidence qualification as stated above. The set DS1 is analyzed with different window shifts. The length of the

training data with 20 ms window shift is half of that with 10 ms window shift, and the length of the training data with 30 ms window shift is one-third.

Table 3. Speaker identification rates using GMM trained with the MFCCs and dPLRM trained with the log power spectrum extracted with different window shifts from three sentences uttered in Session T0/T1/T2 for each sentence.

Training data	Method	Identification rates (%)		
		10 ms shift	20 ms shift	30 ms shift
Word speech	dPLRM	92.7 (89.3,96.0)	91.3 (87.3,94.7)	90.0 (86.0,94.0)
	GMM	89.3 (85.3,93.3)	86.0 (81.3,90.7)	85.3 (80.7,90.0)
Sent. speech	dPLRM	100 (99.3,100)	100 (99.3,100)	100 (99.3,100)
	GMM	99.3 (98.7,100)	98.0 (96.0,99.3)	96.7 (94.0,98.7)

It is interesting to note that the dPLRM method trained on such a coarsely sampled data with 30 ms window shift outperforms the GMM method with full 10 ms shift sampled data. Since the dPLRM can handle nonlinearity more effectively with kernel functions and do discriminating learning interdependently, it is expected to work with a smaller amount of training data.

#### 4. Conclusions

In this paper, speaker identification without outright pre-processing of speech data was shown to be possible by employing the dPLRM. Comparison was made between the dPLRM and GMM-based method in the experiments with training data uttered by 10 male speakers in three sessions, and the dPLRM method with the log power spectrum is competitive with the GMM-based method with the MFCCs. The method outperforms the GMM-based method especially as the amount of training data becomes smaller.

The evaluation of the method with a larger dataset is left for our future study.

#### 5. Acknowledgements

A part of this work was supported by JSPS

Grant-in-Aid for Scientific Research (B) 16300036 and (C) 16500092.

#### 6. References

- [1] <http://www.nist.gov/speech/tests/spk/index.htm>
- [2] S. B. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans. on ASSP*, vol. 28, no. 4, pp. 357-366, 1980.
- [3] D. A. Reynolds, "Speaker Identification and Verification Using Mixture Speaker Models," *Speech Communication*, 17, pp. 91-108, 1995.
- [4] S. S. Stevens, "Psychophysics," John Wiley & Sons, New York, 1975.
- [5] H. A. Murthy, F. Beaufays, L. P. Heck and M. Weintraub, "Robust Text-Independent Speaker Identification over Telephone Channels," *IEEE Trans. on SAP*, vol. 7, no. 5, pp.554-568, 1999.
- [6] A. Biem, S. Katagiri, E. McDermott and B.-H. Juang, "An Application of Discriminative Feature Extraction to Filter-Bank-Based Speech Recognition," *IEEE Trans. SAP*, vol. 9, no. 2, pp. 96-110, 2001.
- [7] C. Miyajima, H. Watanabe, T. Kitamura, and S. Katagiri, "Discriminative feature extraction - Optimization of mel-cepstral features using second-order all-pass warping function," *Proc. Eurospeech*, pp.2-779-782, 1999.
- [8] K. Tanabe, "Penalized Logistic Regression Machines: New methods for statistical prediction 1," *ISM Cooperative Research Report 143*, pp. 163-194, 2001.
- [9] K. Tanabe, "Penalized Logistic Regression Machines: New methods for statistical prediction 2," *Proc. IBIS, Tokyo*, pp. 71-76, 2001.
- [10] K. Tanabe, "Penalized Logistic Regression Machines and Related Linear Numerical Algebra," *KOKYUOKU 1320*, Institute for Mathematical Sciences, Kyoto University, pp. 239-249, 2003.
- [11] T. Matsui and K. Tanabe, "Speaker Identification with dual Penalized Logistic Regression Machine," *Proc. Odyssey*, pp.363-366, Toledo, 2004.
- [12] T. Matsui and K. Tanabe, "Probabilistic Speaker Identification with dual Penalized Logistic Regression Machine," *Proc. ICSLP*, pp. III-1797-1800, 2004.
- [13] <http://htk.eng.cam.ac.uk>, the hidden Markov model toolkit (HTK).

# 応用：工学分野等における逆問題の展開

大阪大学大学院工学研究科 久保司郎

## 1 はじめに

工学，理学，医学などの多くの分野で応答や結果などの出力から，原因や入力を推定する逆問題 (inverse problem) の重要性が認識され，盛んに取り扱われるようになってきている [1]–[32]．逆問題の取扱いを逆問題解析あるいは逆解析 (inverse analysis) という．

本稿では，逆問題とそれを取り扱う逆解析法の特徴と適用例の紹介を行う．

## 2 逆問題の意義と逆解析の特性

### 2.1 逆問題の意義

逆問題 (inverse problem) の代表例として X 線 CT (断層法) があげられる．これは，種々の方向から X 線を照射したときに得られる投影像から，X 線吸収能の分布を表す断面像を構築するものである．このように直接は求められないものを推定することが多くの分野で要求されおり [1]–[32]，構造物の状態や特性を推定することが必要になることがある．たとえば，建造されてから何年も使用されてきたプラントや構造・機器に対しては，経年劣化の程度を評価し，その部材をとのまま使用するか，補修するか，取り替えるかといった判断が求められる．このためたとえば，プラントや構造・機器に発生し，存在する欠陥やき裂を検出し，その大きさと形状を推定することが求められる．また材料特性値の分布を推定することができれば，異常な箇所を検出したり，製品が設計どおりに製作されているかを調べることができる．また，構造の強度に大きな影響を与える残留応力の分布を推定することも必要とされている．これらの評価により長期間使用された機器や部材に対して，経済性を損なうことなく安全性を確保することができる．また，所要の特性や性能を与えるような機器の形状や材料特性の分布を推定することは，設計の問題に直結する．

これらのような多くの逆問題がある．以下においては，工学，理学，医学等の分野における逆問題とその取扱いを考慮しながら，逆問題の分類と特性について述べる．

### 2.2 逆問題の定義と分類

これまで扱われてきた逆問題には多様なものが存在し，研究分野によって逆問題の意味するところが異なっている．そこでまず，種々の分野で現れる逆問題を総括して考えられるような，逆問題の定義について論じておく．

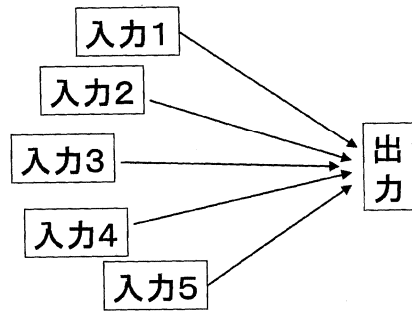


図 1: 多数入力と出力

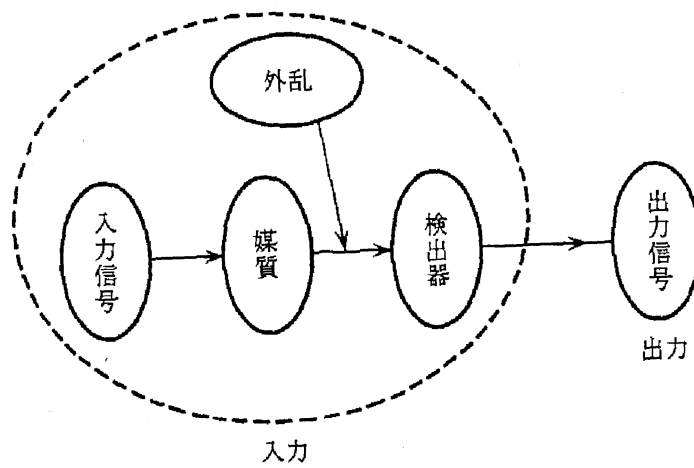
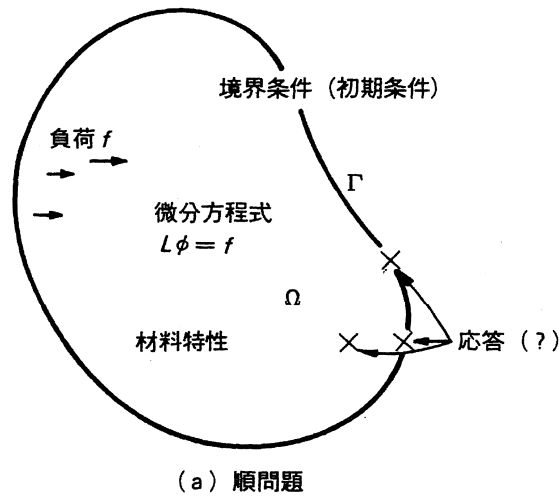


図 2: 信号伝達システムを含めた多数入力と出力



- (a) 領域Ωと境界Γ  
 (b) 場の支配方程式  
 (c) 境界条件 (初期条件) → 応答  
 (d) 負荷  $f$  (Γ上の未知境界値)  
 (e) 材料特性 (Ω内の物理量φ)
- (b) 順解析

図 3: 順問題と順解析

逆問題は順問題以外のものであると考えれば、順問題 (direct problem) がどのような入力から構成されるかが明らかであれば、逆問題の明確な定義が可能である。図 1 のように、一般には入力は 1 個ではなく複数個ある。すべての入力を与えられれば、順方向の手順により出力が決まる。図 2 のように、信号を伝達するシステムも入力の一つとして捉えることができる。入力の中に未知のものがあるときに、出力をもとに未知入力を推定するものが逆問題である。入力の中のいずれが未知であるかによって異なった問題が提起されることになるが、それらはすべて逆問題である。

このように順問題が何であるかに従い、逆問題としてどのようなものが存在するかが決まり、逆問題の分類ができる。

一つの分類の例として、この考え方にしがたい、物理量などの空間変化および時間変化を扱う場の解析に関連した逆問題について論じる [6],[17]。現象あるいは状態を表現している何らかの量を  $\phi$  で表すとき、 $\phi$  の空間変化あるいは時間変化に関する支配微分方程式

$$L(\kappa)\phi = f \quad (1)$$

を用いて、 $\phi$  の応答を求める順解析 (direct analysis) が行われる。ここに、 $L(\kappa)$  は材料特性  $\kappa$  を含んだ作用素、 $f$  は領域内の負荷に相当する項である。図 3(a) および (b) のように、この順解析を実施するためには、次のような情報、すなわち入力が不可欠である。

- (a) 対象とする領域  $\Omega$  とその境界  $\Gamma$  の位置と形状
- (b)  $\phi$  に関する場の支配方程式 ( $L\phi = f$ )
- (c)  $\phi$  あるいはその微係数に関する境界条件 (もし必要があれば初期条件も含む)

- (d) 領域内の負荷  $f$
- (e) 支配方程式に含まれる材料特性  $\kappa$  の分布

もし、これらの要件のうちのいずれかが欠落しておれば、順解析を行うことはできない。上述のように、逆問題は順問題以外のものとする。このとき、場の解析では (a) から (e) までのいずれか、あるいは二つ以上の組合せに未知のものがあるときに、これらを同定しようとするものが逆問題であり、これを扱う解析が逆解析 (inverse analysis) あるいは逆問題解析と呼ばれるものである。上記の (a) から (e) と対応する逆問題としては、図 4(a) のように、次のものがある。

- (A) 領域  $\Omega$  の形状あるいは領域内の未知境界を同定する問題 (領域／境界逆問題)
- (B)  $\phi$  の場を支配している微分方程式を求める問題 (支配方程式逆問題)
- (C) 境界  $\Gamma$  の一部あるいは全部における境界値、あるいは領域  $\Omega$  内で定義される初期値を求める問題 (境界値／初期値逆問題)
- (D) 領域  $\Omega$  の内部で作用している負荷  $f$  を求める問題 (負荷逆問題)
- (E) 領域  $\Omega$  の内部における材料特性  $\kappa$  を推定する問題 (材料特性逆問題)

これらの逆問題の組合せも逆問題である。逆解析の操作の呼称としては、問題に応じて同定、推定、再構成、逆変換、逆計算、逆設計などが用いられる。

場の解析以外の問題についても、順問題が与えられれば同様に逆問題の分類ができる。多くの場合に対し、上記の分類があてはまるので、以下の議論ではこの分類を参照することとする。

逆問題解析では順問題解析に比べて入力不足している。逆問題を解くためには、何らかの追加情報が必要である。この情報として利用できる主なものは、図 4(b) に示すように、順解析で解析の対象とした出力、すなわち応答である。この情報を一次情報と呼ぶことにする。一次情報としては、領域内や境界上で得られた  $\phi$ 、その微係数などが考えられ、これらは計測などにより獲得することができる。また、逆設計においては一次情報は設計条件として与えられる。

境界上で与えられる境界一次情報を用いる場合には、応答が与えられる部分の境界条件は、順問題の境界条件に比べ多い。このような境界条件を過剰規定境界条件、その境界を過剰規定境界と呼ぶ。過剰規定境界を設けることができない場合に対しては、領域の内部において得られる領域内一次情報を用いることが考えられる。

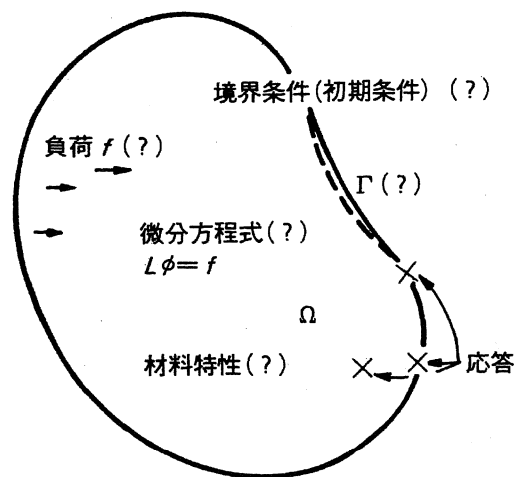
応答以外で逆解析に利用できる情報としては、解の上限、下限などの知見、経験などによる先験情報や物理的な制約条件がある。この情報には定性的なものも多いが、逆問題解析には重要な役割を果たす。この情報を、二次情報あるいは補助情報と呼ぶ。

解やその候補が、二次情報を満足するような場合には、それが許容であると呼ぶことにする。二次情報を用いると、解の存在範囲を狭めることができる。

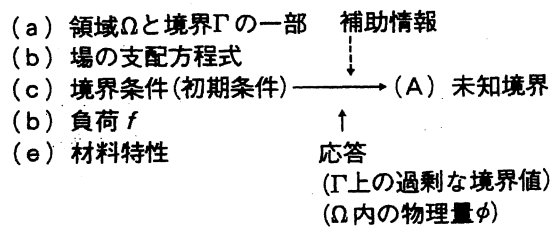
## 2.3 逆問題の不適切性

一般に提起された問題が適切である (well-posed) こととは、(1) 解の存在性、(2) 解の一意性および (3) 解の連続性あるいは安定性の三つの要件がすべて満足されていることを表している。これらの要件のうち少なくとも一つが失われている問題は、不適切である





(a) 逆問題



(b) 逆解析の例(領域/境界逆問題)

図 4: 逆問題と逆解析

(ill-posed) という。順問題とは異なり、通常の場合、逆問題はこのような不適切性を有している。

適切性の三つの要件のうち、解の存在性と一意性がともに成立しているような場合でも、(3)の安定性が失われており、問題が不適切となっていることが多い。たとえば、ラプラスの微分方程式のもとでは変数  $\phi$  の微小な変動があっても、その影響は変動がある部分から離れるにつれ急速に減衰する<sup>[33]</sup>。

また、サン・ブナンの原理より、弾性体の変形の問題においても釣合いを満足する力の系が作用してもその影響は空間的に急速に弱まる。逆問題においては、同定の対象とする部分より離れた位置で得られた観測値などを一次情報として用いることが多いため、減衰した観測情報からもとの状態を同定することになる。このため、数値計算上は解くべき方程式が悪条件になり、結果として一次情報に含まれる誤差が著しく拡大されて、推定結果に悪い影響を及ぼす。

以上のように、多種多様な逆問題が存在する。以下では、工学分野等における逆問題とその取扱いの例について述べる。

### 3 領域／境界逆問題とその取扱いの例

#### 3.1 能動型および受動型電気ポテンシャル CT 法によるき裂の同定

構造物や部材の安全性を評価するためには、その中に存在する欠陥やき裂を非破壊的に測定することが必要になる。欠陥やき裂の非破壊計測は、何らかの応答をもとにき裂や欠陥を同定するものである。領域／境界逆問題のひとつである。き裂の測定には、物体に能動的に通電したときの応答、すなわち電気ポテンシャル分布を用いる電気ポテンシャル法やインピーダンス・プレチスモグラフィがよく用いられている<sup>[34]–[38]</sup>。ここでは、電気ポテンシャル CT 法<sup>[39]–[45]</sup>を中心に、電気ポテンシャル情報からのき裂同定について述べる。

直流電流のもとでは、電気ポテンシャル分布は、次式により支配されている。

$$\nabla(\lambda \nabla \phi) = 0 \quad (2)$$

ここで、 $\lambda$  は電気伝導率である。欠陥やき裂の位置、形状が与えられれば、境界要素法や有限要素法などを用いて、この方程式のもとにおける電気ポテンシャル分布を解析することができる。

電気ポテンシャル情報からのき裂同定の一意性については理論的検討がなされている<sup>[46]–[47]</sup>。まず、き裂存在面が既知である場合について考える。き裂を有する物体に通電し、そのときに物体表面の一部において電気ポテンシャル  $\phi$  とその法線方向勾配である流束  $q$  の分布が厳密に与えられているものとする。このとき、一組の  $\phi$  と  $q$  の分布の情報より、き裂存在面内のき裂の位置および形状が一意に定められることが、境界値逆問題の解の一意性を用いて示される。き裂存在面の位置が未知である場合には、 $n$  次元物体内の単一き裂の同定には  $n$  組の  $\phi$  と  $q$  の分布が必要である。ただし、一意性は保証されても安定性がないため、この問題も不適切である。

このき裂同定に関する逆問題の解析に、二つの解析手法が適用されている。その一つ

は、未知境界同定法と呼ばれている。この方法の原理は、き裂を含む面  $\Gamma_0$  が既知である場合に境界積分方程式を未知境界値について解いて、 $\Gamma_0$  上の流束  $q$  の分布を推定し、 $q$  が0となる部分よりき裂を同定することにある。したがって、この方法ではき裂という未知境界の同定問題が境界値の同定問題に置き換えられている。

き裂同定に対するもう一つの逆解析手法は、境界残差最小化法と呼ばれており、仮定き裂に対する境界上の電気ポテンシャル分布の計算結果と測定結果の間の残差  $R$ （式(25)または(26)）より、準解を求めるものである。すなわち、この方法では、適当なルールにしたがってき裂位置・形状を仮定し、仮定されたき裂に対し計算により求めた電気ポテンシャル分布  $\phi^{(c)}$  を測定されたもの  $\phi^{(m)}$  と比較し、両者が最もよく一致するき裂を探索により求める。経験などの先験情報に基づいて、仮定するき裂の候補を、絞ることができる。

この境界残差最小化法を用いて、未知面内に存在する2次元斜めき裂の位置と形状を推定する実験が行われている。電気ポテンシャルの計算値  $\phi^{(c)}$  は境界要素法により求められている。上記のように、単一の電気ポテンシャル分布のみでは未知面内にあるき裂を一意に同定することはできない。そこで、電流の負荷として複数のものを与えそれぞれに対して試験片外周の電気ポテンシャル分布を測定し、その結果をき裂同定に用いる、マルチ電流負荷方式が提案されている<sup>[41]</sup>。この実験の結果によれば、単一の電流負荷に対する電気ポテンシャル分布のみを用いた場合には実際のき裂とかけはなれたき裂が同定されることがあるのに対して、マルチ電流負荷方式を採用した場合には、良好な同定結果が得られている。

図5のような3次元表面き裂および内部き裂が、境界残差最小化法を用いて推定されている<sup>[42]–[43]</sup>。この解析では、き裂位置・形状の探索を効率化するため、3次元逆解析の前処理として2次元近似走査解析を用い、また2次元および3次元解析のそれぞれについてき裂の仮定の最小単位を段階的に細かくする階層化探索法が採用されている。電気ポテンシャルの計測システムの例を図6に示す。ポテンシャルの計測値を用いて、表面き裂を背面のポテンシャル分布あるいは開口側前面および背面より推定した逆解析の結果が、図7にハッチング部で示されている<sup>[42]</sup>。図の破線は放電加工により導入された実き裂の形状である。図のように良好な推定結果が得られている。電気ポテンシャル分布の計算には少ない自由度で高精度の解が得られる境界要素法が用いられている。電気ポテンシャル分布の解析結果をデータベース化しておくことにより、効率的に解析を行うことができる<sup>[44]</sup>。このほか、感度解析を用いた評価をき裂形状同定に適用したもの<sup>[48]</sup>がある。

電気伝導率  $\lambda$  が場所により異なる異種接合体内に存在する界面欠陥の同定に対しても、電気ポテンシャルCT法が適用されている<sup>[49]–[51]</sup>。電気伝導率  $\lambda$  に異方性がある問題に対する適用性も調べられている<sup>[52]</sup>。また、直流電気ポテンシャル法の感度を向上させるため、近接端子を用いる方法が提案されている<sup>[53]–[54]</sup>。

上記の電気ポテンシャルCT法は、通電を必要とするいわゆる能動型（アクティブ）検査法である。圧電効果を持つピエゾフィルムを物体表面に貼り付ければ、物体に力が加わったときの物体のひずみ分布に応じたピエゾフィルム上の電気ポテンシャル分布を、外部から通電することなく、受動的に得ることができる。き裂があるときのひずみ分布はき裂の影響を受け、ピエゾフィルム上の電気ポテンシャル分布はき裂の位置と寸法を反映したものとなる。したがってピエゾフィルム上の電気ポテンシャル分布をもとにき裂を逆問

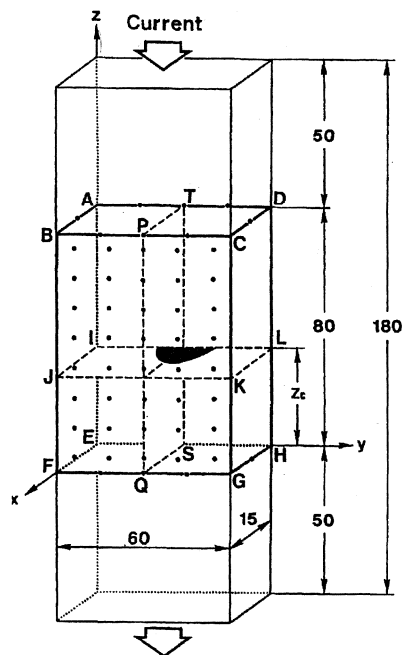


図 5: 表面き裂を有する試験片

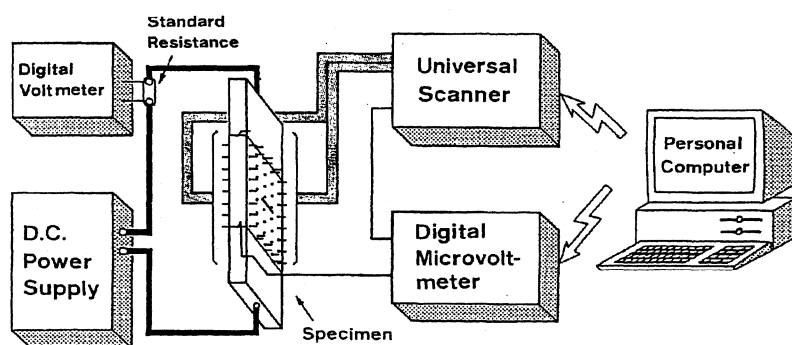


図 6: 能動型電気ポテンシャル CT 法によるき裂同定のための電気ポテンシャル分布測定システム

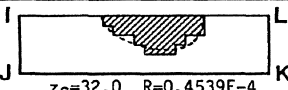
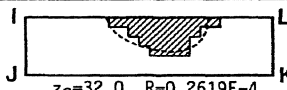
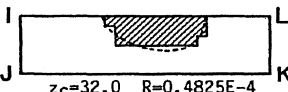
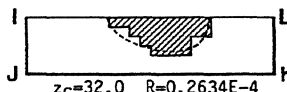
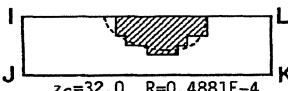
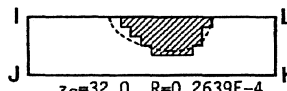
Cracks Estimated by Electric Potential CT Method	
(i) Using Front and Back Side Potential Data	(ii) Using Back Side Potential Data
 <p><math>z_c=32.0</math> <math>R=0.4539E-4</math> <math>S_c=0.188</math></p>	 <p><math>z_c=32.0</math> <math>R=0.2619E-4</math> <math>S_c=0.188</math></p>
 <p><math>z_c=32.0</math> <math>R=0.4825E-4</math> <math>S_c=0.188</math></p>	 <p><math>z_c=32.0</math> <math>R=0.2634E-4</math> <math>S_c=0.188</math></p>
 <p><math>z_c=32.0</math> <math>R=0.4881E-4</math> <math>S_c=0.188</math></p>	 <p><math>z_c=32.0</math> <math>R=0.2639E-4</math> <math>S_c=0.181</math></p>

図 7: 能動型電気ポテンシャル CT 法による表面き裂の同定結果

題的に同定することができる [55]–[57]。このような受動的な手法により、適用性の幅が広がることが考えられる。

一般の弾性体に対しては、応力各成分からなる応力ベクトル  $\{\sigma\}$  と、ひずみ各成分からなるひずみベクトル  $\{\epsilon\}$  の間に、

$$\{\sigma\} = [C]\{\epsilon\}$$

なる関係が成立する。ここに、 $[C]$  は材料剛性マトリックスである。これに対し、 piezoelectric material では、ひずみと電場の間に連成関係があり、 $\{\sigma\}$  は、 $\{\epsilon\}$  とともにポテンシャルの勾配で表される電界ベクトル  $\{E\}$  の影響を受け、逆に電気変位ベクトル  $\{D\}$  は  $\{E\}$  とともに  $\{\epsilon\}$  の影響を受ける。

$$\{\sigma\} = [C]\{\epsilon\} - [e]^T\{E\}$$

$$\{D\} = [e]\{\epsilon\} + [g]\{E\}$$

ここに、 $[e]$  および  $[g]$  は圧電定数応力マトリックスおよび誘電マトリックスである。

圧電体では、電気ポテンシャル場と応力-ひずみ場が連成しており、電気ポテンシャル場、応力-ひずみ場という異なる場の間のエネルギー変換を伴うため、エネルギー関数の形で書かれる汎関数の適用を基本とする有限要素法を適用すると、有限要素方程式が得られる。また境界要素法を用いて電気ポテンシャル分布を計算することができる。

二次元貫通き裂に対する受動型電気ポテンシャル CT 法の適用性が調べられている [57]。まず図 8 のように弾性材料に piezoelectric film を貼り付けたものが用いられた。図のような  $x$  軸に垂直な二次元貫通き裂を導入されている。き裂の位置、寸法を表わすき裂パラメータは、き裂半長  $a$ 、き裂の物体表面からの深度  $h$ 、き裂の  $x$  方向の位置  $x_c$  である。

図 9 は、き裂深さ  $h$  が piezoelectric film 上の電気ポテンシャル分布に与える影響を調べたものである。電気ポテンシャル分布の計測例を図 10 に示す。図のように、計測されたポ

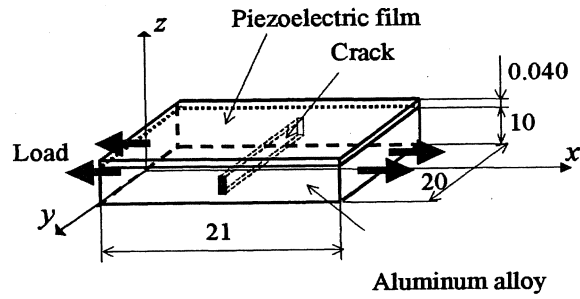


図 8: 受動型電気ポテンシャル CT 法による貫通き裂の同定

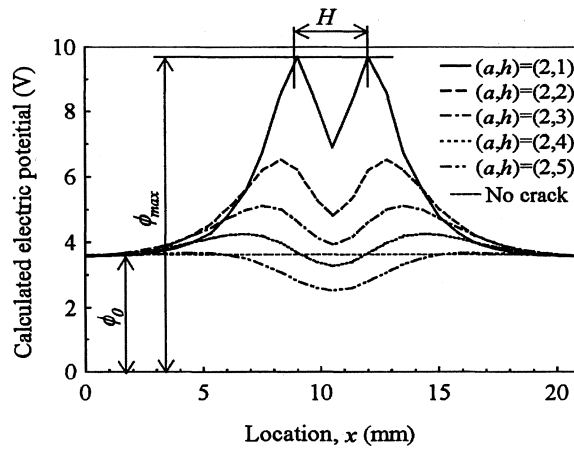


図 9: ピエゾフィルム上の電気ポテンシャル分布に及ぼすき裂深さ  $h$  の影響

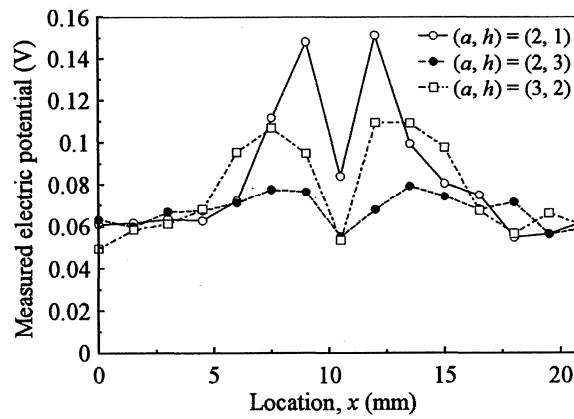


図 10: ピエゾフィルム上の電気ポテンシャル分布の計測結果

表 1: 受動型電気ポテンシャル CT 法によるき裂の同定結果

$(a, h)$		Parameters of crack			Residual $R_s$
		$a$	$h$	$x_c$	
(2, 1)	Estimated	2.15	1.25	10.5	0.24
	Actual	2.0	1.0	10.5	
	Error (%)	7.5	25	0	
(3, 2)	Estimated	3.10	2.18	10.3	0.32
	Actual	3.0	2.0	10.5	
	Error (%)	3.3	9.0	1.9	
(2, 3)	Estimated	1.93	3.25	10.8	0.076
	Actual	2.0	3.0	10.5	
	Error (%)	3.5	8.3	2.9	

テンシャル分布は計算と同様の分布形を示しており、この分布からき裂の位置および寸法を推定することができる。き裂同定結果を表 1 に示す。表のようにき裂が良好に同定されている。

複数のき裂や 3 次元物体中の表面き裂についても、受動型電気ポテンシャル CT 法によりき裂の個数、位置および形状が推定できることが明らかにされている [58]–[59]。

### 3.2 渦電流探傷による欠陥形状の定量評価

渦電流探傷は、欠陥やき裂の検出によく用いられている。渦電流探傷は、検出感度が高いため欠陥やき裂の検出には向いているが、これを定量測定と結び付けるのは容易ではない。そこで、電磁場の数値解析結果を用いることにより、欠陥やき裂の定量的同定が行われている [60]–[62]。

### 3.3 動特性あるいは熱的応答を用いた欠陥同定

物体内に存在する欠陥を同定するために、物体に調和加振力を与えたときの動的な応答を利用する方法が検討されている [63]。この方法でも、境界要素法による応答に関する計算値と計測値の残差を調べ、残差が小さくなるような欠陥が同定結果として採用されている。単一の調和加振力を加えたときの固有角振動数を用いて空洞の形状を推定した結果では、き裂の大きさはわかっていても形状は同定できないが、種々の加振状態に対する結果を総合した場合には、よい結果が得られる [64]。複数の欠陥があるときの等価だ円欠陥が同定されている [64]。複数欠陥の同定には、位相の同定に有望な遺伝的アルゴリズムが適用されている [65]–[66]。

欠陥を有する物体に熱負荷を与えたときには、その熱的応答に変化が出る。そこで、表面温度分布をサーモグラフィなどにより求め、これを計算結果と比較することにより、欠陥を定量的に同定する試みがなされており、サーマルトモグラフィ、サーマルウェーブ CT などの名称がつけられている [67]–[70]。

### 3.4 波動応答を用いた欠陥同定

欠陥等が存在すると弾性波の散乱が起こることを利用して、欠陥を同定する方法は、逆散乱問題、超音波探傷として知られている<sup>[71]–[76]</sup>。異種材料の界面に存在する欠陥を推定するためにも、この方法が適用されている。

波動場  $u$  は、次のヘルムホルツ方程式により支配されている。

$$\nabla^2 u + k^2 u = 0 \quad (3)$$

ここに、 $k$  は波数である。欠陥等の位置が既知であるときには、このヘルムホルツ方程式を解く順解析を行うことにより、波動場が求められる。欠陥部で1、他の部分で0となるような特性関数を導入し、その値を散乱波より求めることにより、欠陥の位置と形状の推定が行われている。

音響診断に対しても、階層型ニューラルネットワークが適用されている<sup>[77]–[79]</sup>。

## 4 境界値逆問題とその取扱いの例

### 4.1 ポテンシャルに関する境界値逆問題

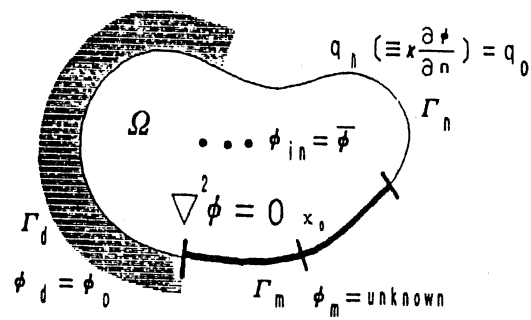
境界値逆問題の例として、図 11(a) のようなガルバニック腐食逆問題について述べる。ガルバニック腐食とは、異種金属が電氣的に接触した状態で腐食性溶液などの中に置かれると、一方がアノードに、他方がカソードになり、それらの電気ポテンシャル差により電気が流れるものである。このときの電気ポテンシャルの分布は、各金属の分極特性、すなわち電気ポテンシャルと電流密度の関数に依存する。金属の分極特性は未知であることが多く、そのうえ金属表面の電気ポテンシャルおよび電流密度の分布は測定することは通常困難である。このため、溶液内部の電気ポテンシャルの測定値を用いて境界上の電気ポテンシャルと電流密度を推定する逆問題が提起されている<sup>[80]–[82]</sup>。

この問題に対して、内部の測定点の電気ポテンシャルに関する境界積分方程式が導入されている。この式を未知境界値について解くことができる。測定点の位置が不完全規定境界部から離れているときには、振動する不安定な解が得られる。特異値分解を適用して境界値を求めると、電流密度の推定結果は正解に近いが、本来は二金属の接触部で不連続となるべき分布が、平滑化されてしまう。チホノフの適切化法に準じて曲率に対応した二階微分が小さくなるように処罰項を付加しても同様の結果が得られている。

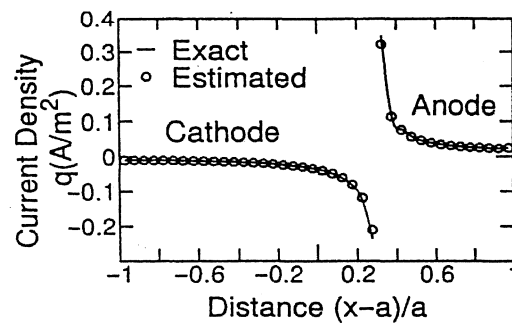
そこで、二種類の金属のみが溶液と接していることを用いて、金属の分極特性や継ぎ目の位置に関するあいまいな情報を用いて、逆解析値をもとに不完全規定境界部をふたつの金属に対応する境界に分離することが行われている。この結果を用いると、不完全規定境界上の境界値は最終的には順解析より求められる。その結果、図 11(b) に示すような良好な推定値が、安定して求められる<sup>[82]</sup>。

ポテンシャルに関する境界値逆問題の構造を数理的に解明し、この結果をもとに有効な適切化手法を構築する試みがなされている<sup>[83]–[84]</sup>。図 12 のように、領域内で定義された関数  $\phi$  がラプラス方程式を満たすとき、境界 AD で境界値が与えられておらず、これに対して境界 BC では境界値として  $\phi$  だけでなく、その法線方向勾配である流束  $q$  が与えられ





(a) 解析対象と境界条件



(b) 境界値の推定結果

図 11: ガルバニック腐食における境界値逆問題

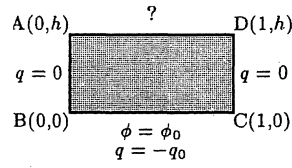


図 12: ポテンシャルに関する境界値逆問題

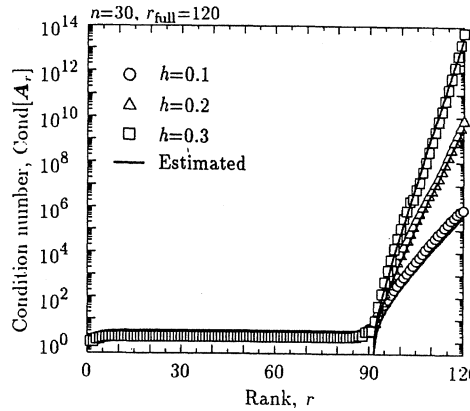


図 13: ランクを低下させたときの条件数の挙動に関する推定結果と計算結果の比較

ている問題が取り扱われている。この過剰な境界値を用いて AD 上の不完全規定境界値を推定する問題の数値的構造が数理的に調べられ、観測値に含まれる誤差が推定結果に及ぼす拡大率を表す条件数の挙動が求められている。特異値分解を適用しランク低下法による適切化を行ったときの条件数の挙動が推定されている。図 13 に示されているように、推定結果と数値解析結果がよく一致している<sup>[83]</sup>。図 14 にポテンシャルに関する境界値の推定例が示されている。許容条件数法などにより推定されたランクを用いることにより、適切化が達成されていることがわかる。

交替境界要素逆解法<sup>[85]–[86]</sup>を用いた逆解析も行われている<sup>[87]–[88]</sup>。この解析では、過剰規定境界値のうちポテンシャル  $\phi$  に関するものを用いた順解析と流束  $q$  に関するものを用いた順解析を、交互に繰り返すことにより、境界値の推定が行われている。観測値に含まれる誤差を考慮するため、繰返しの打ち切り数を適切化パラメータとし、これを食違い量基準により推定することが行われている。

## 4.2 表面力および変位の推定問題

ふたつの物体が接触している境界上の接触力と変位を、その状態を乱すことなく測定することは困難である。そこで、接触部の状態を他の部分で測定される変位と表面力を用い

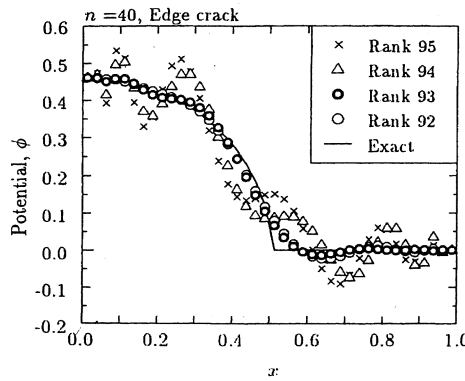


図 14: ポテンシャル分布に関する境界値の推定結果

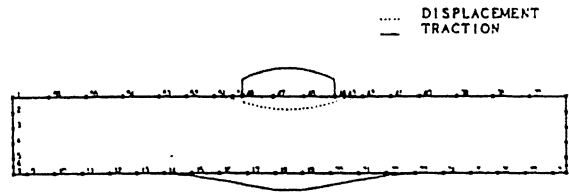


図 15: 接触部の接触力および変位の推定

て推定する方法が考えられている<sup>[89]–[91]</sup>。この方法では、有限要素法により境界を離散化した有限要素関係式が用いられている。この式に測定値を代入することにより導かれた方程式が、擬似逆行列を用いて未知境界値について解かれている。実測により得られた下面の圧力のデータをもとに、接触力と変位を求めた結果の例を図 15 に示す<sup>[90]</sup>。なお、接触領域が未知である場合に対しては、力の二乗和が最小となる条件により接触領域の大きさが推定されている。このことは、チホノフの適切化法において、関数のノルムの二乗が処罰項としてよく採用されることと関連させて考えることができる。

また、接触力同定問題に対し適切化をはかるため、関数展開法が適用されている<sup>[92]</sup>。この方法では有限要素関係式が未知境界値を表す関数展開の係数により表現され、これに観測値を代入した式が解かれている。この方法の適切化パラメータは関数展開の項数である。その最適値は、AIC 最小化法または推定誤差基準により求められる。このような関数展開法では解空間に制約を加えていることになるが、同様のことは境界上にキーノードを導入し、それらに対する表面力を用いて表面力分布を表現する方法によっても実現できる<sup>[93]</sup>。界面における変位の食い違い量と界面力とを積分方程式を用いて求める方法も検討されており、解の一意性も議論されている<sup>[94]</sup>。適切化にはチホノフの方法が適用されている。

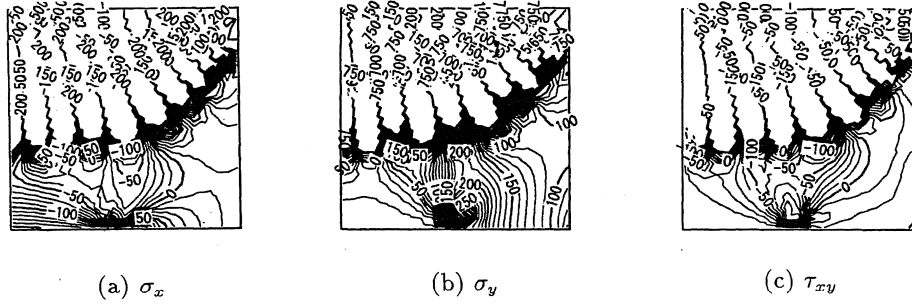


図 16: 境界値推定を媒介とした主応力和からの応力成分の推定（適切化を行わない場合）

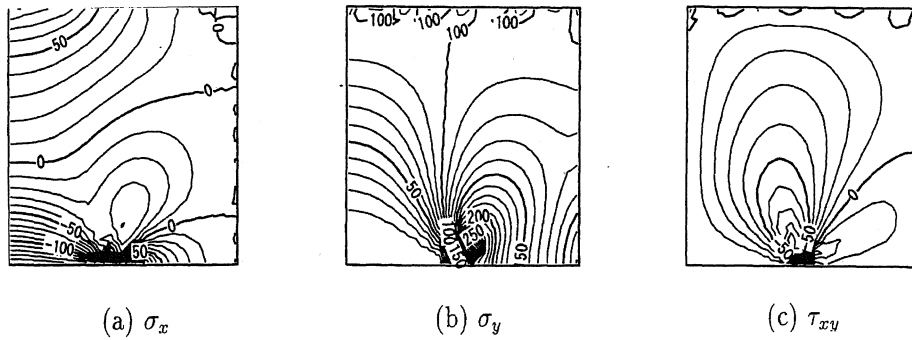


図 17: 境界値推定を媒介とした主応力和からの応力成分の推定（適切化を行った場合）

静弾性境界値逆問題の構造を数理的に解明し、ランク低下に伴う条件数の挙動が解析的に予測され、その結果をもとに有効な適切化手法を構築する試みがなされている<sup>[95]</sup>。

定常動弾性問題について、境界上の負荷をいくつかの観測点における動的応答から推定する方法が調べられている<sup>[96]</sup>。

#### 4.3 境界値推定を媒介とした応力およびひずみ分布の推定

物体の熱弾性効果により、応力の第一不変量である主応力和が  $\Delta s$  だけ変化すると温度  $T$  が

$$\Delta T = -K_m T \Delta s \quad (4)$$

だけ変化する。熱弾性応力測定法は、この効果に基づき、温度変化から主応力和の変化を求めるものである。ここに、 $K_m$  は熱弾性定数である。この式より、物体の温度変化の画像から、主応力の変化を求めることが可能であるが、応力の各成分を直接求めることはで

きない。そこで、逆解析を用いて、境界値逆解析を媒介として、物体に作用している応力の成分分離を行うことが行われている<sup>[97]–[99]</sup>。この主応力の分離の解析は、以下の2段階に分けて行われている。

(1) 主応力和のデータから境界値を推定する逆問題

(2) 境界値に基づいて各応力成分を求める順問題

適切化のため、特異値分解を用いたランク低下法などが適用されている。また、種類の異なる情報として、ひずみゲージの読みをあわせて用いると、同定の精度が上がることを示されている<sup>[100]</sup>。この問題では、主応力がつねに零となるせん断型のモードが同定できないことについても、議論がなされている。

主応力和からの各応力成分の推定において、適切化を行わない場合と行った場合の数値シミュレーション結果を、それぞれ図 16 および 17 に示す<sup>[99]</sup>。図のように、適切化を行わない場合には、誤差が大きく拡大され、応力分布に不自然な変動が現れるが、適切化を行うことにより、正解に近い推定結果が得られる。

応力の成分分離と関連して、平面内で応力とひずみの場が測定されているとき、測定値にはノイズが入っているため、これらは一般に連続体の基礎関係式である適合条件式と釣合式を満足しない。そこで、適合条件式と釣合式が満足されるように、ペナルティ項を導入することにより、測定された場のフィルタリングが行われている<sup>[101]</sup>。その結果、ノイズを含む測定結果の誤差を低減できることが示されている。また、物体の表面のひずみの観測データから、弾性ひずみと塑性ひずみとを直接求めることはできない。そこで、ひずみの測定結果に増分塑性の構成則をあてはめることにより、応力を求め、ひずみを弾性成分と塑性成分に分離することが行われている<sup>[102]</sup>。

## 4.4 心電図逆問題

体表面で計測した電気ポテンシャル分布から、心臓の表面における電気ポテンシャル分布の時間的変化を推定する問題が取り扱われている<sup>[103]–[104]</sup>。この問題では、体表面では流束が0であるので、体表面における過剰な境界値を用いて、境界値がまったく与えられていない心臓表面における境界値を推定することになる。

## 5 負荷逆問題とその取扱いの例

### 5.1 残留応力および残留ひずみの同定

構造物やその要素が製造過程中にうけた残留ひずみが適合性を満足しない場合には、残留応力が生じる。この残留応力は強度に大きな影響を与えるため、残留応力の測定に関して多くの検討結果がある<sup>[105]–[112]</sup>。切断時のひずみ変化のような間接的な測定値より、残留応力および残留ひずみを推定する問題を領域内の負荷に関する一種の逆問題と解釈することができる。

残留ひずみの推定において重要なことは、一般に解の一意性が成立しないことである。このことは、適合条件を満足するような残留ひずみ分布がある場合には、そのことが測定

結果に何らの影響も与えないことからわかる。また、残留応力は自己平衡性を満足しているものでなければならない。

3次元残留応力の分布を推定する方法として、自由膨張や収縮に対応する残留応力に無関係な固有ひずみを除いた有効固有ひずみを、切断ときに計測されるひずみより求める方法が提案されている<sup>[105]</sup>。また、計測点を除去したときのマトリックスの性質を調べることで、良好な計測点を選定する方法が提案されている。

残留応力分布を直接推定する場合には、残留応力分布が自己平衡型であることを積極的に利用することができる。部材にき裂が入ったことにより再分布した残留応力より、き裂が入っていない初期の残留応力分布を求める問題にこの考えが適用されている<sup>[113]–[114]</sup>。この方法では、まず力に関するいくつかの平衡条件を満足する基本分布関数を導入し、求めるべき残留応力分布をこの基本分布関数の重ね合せにより表現する。このことにより比較的少数個のパラメータを用いて残留応力分布を効率的に表現することができる。またこの方法では、測定値が推定結果に及ぼす影響を評価する逆感度マトリックスが提案されている。逆感度マトリックスを評価することにより、推定結果に含まれる誤差が予測でき、測定点に関するよい組合せを選定することができる。このように基本分布関数を導入する方法も、解空間を制約することにより適切化を行っているものと解釈できる。

固有ひずみの空間変化により生まれる不適合度を媒介として、固有ひずみの分布を推定する方法が提案され、解の一意性を回復するための条件が議論されている<sup>[115]</sup>。また、特異値分解と先験情報を用いて固有ひずみ分布を少数の変位計測から求める方法が提案されている<sup>[116]</sup>。

## 5.2 AE および地震波の原波形解析

破壊の状況を調べる方法に、材料や構造物何われが発生したとき生じる弾性波を検出するアコースティック・エミッション(AE)がある。AE法で、割れ発生位置とどのような割れが生じたかを表すAE原波形 $D(t)$ の形態を求めようとするものが、AE原波形解析<sup>[117]</sup>である。地震波の観測でも地震発生源と発生後の割れの進展が重要である。しかし、図18のように原波形は変換子に到達するまでに媒体のグリーン関数 $G(t)$ により変化を受け、さらに計測系の応答関数 $S(t)$ により変換させられている。この結果、検出波形は $V(t)$ は、 $D(t)$ 、 $G(t)$ および $S(t)$ のたたみ込み積分として $D(t)*G(t)*S(t)$ で与えられる。したがって、Fourier変換および逆Fourier変換をそれぞれ $F$ および $F^{-1}$ で表すと、原波形は

$$D(t) = F^{-1}[F[V(t)]/F[G(t)]F[S(t)]] \quad (5)$$

により計算できる。グリーン関数 $G(t)$ は、解析解や数値解により与えられる。応答関数 $S(t)$ は破断特性が既知のシャープペンシルの芯を、媒体のグリーン関数が既知の媒体の端で折ったときの検出波形から求めることができる。図19は、検出されたAE波形と、検出波形に上記の処理を施すことにより得られた原波形の例である。

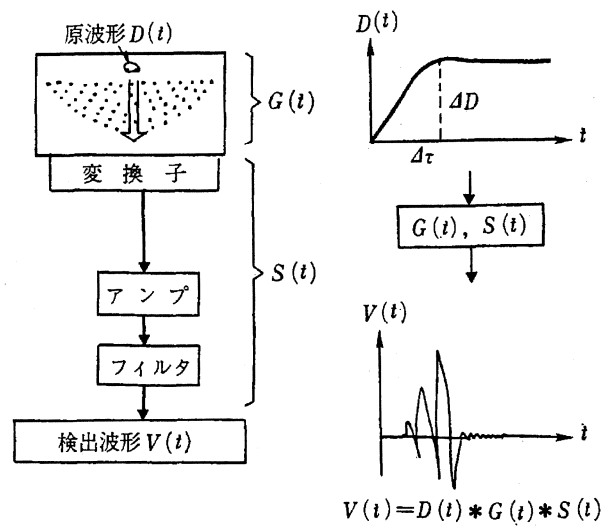


図 18: AE 原波形と検出波形

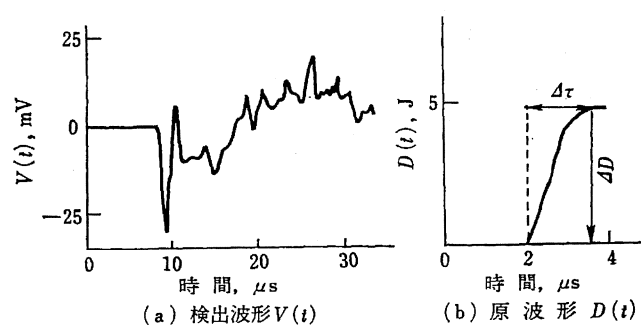


図 19: 微視割れに対する AE 原波形の解析結果

### 5.3 領域内の負荷の同定

状態量  $\phi$  の変化を支配する式 (1) において  $f$  は領域内の熱負荷や物体力などを表す負荷項である。この負荷を同定する方法として、随伴境界積分法が提案されている。この方法では、領域  $\Omega$  内で定義される適当な関数  $\psi^*$  を導入するとき、次の恒等関係式が成り立つことを利用するものである。

$$\int_{\Omega} (\psi L\phi - \phi L^*\psi) d\Omega = \int_{\Gamma} b(\phi, \psi) d\Gamma \quad (6)$$

ここに、 $L^*$  は随伴作用素、 $\int_{\Gamma} b(\phi, \psi) d\Gamma$  は  $\phi$  および  $\psi$  により定義される境界積分である。 $\psi$  として領域内で

$$L^*\psi = 0 \quad (7)$$

を満足するものを採用すると、これらの式および式 (1) より、

$$\int_{\Omega} \psi f d\Omega = \int_{\Gamma} b(\phi, \psi) d\Gamma \quad (8)$$

が得られる。この式は、領域  $\Omega$  内の負荷  $f$  に関する情報が、式 (?) を満足するように選ばれた種々の  $\psi$  に対して評価される境界積分から得られることを表している。

物体中の熱源や力の位置と大きさを推定する数値シミュレーションが行われている [118]–[120]。

## 6 材料特性逆問題とその取扱いの例

### 6.1 多層不均質弾性体中の弾性波の到達速度からの伝播速度分布の推定

不均質弾性体の内部構造を推定するため、表面上の2点間の弾性波の到達速度を用いるものがある [121]。表面上の弾性波の発生源より  $x$  離れた位置における点までの波の到達時間を表す走時曲線から深さに伴う伝播速度の変化が推定されている。

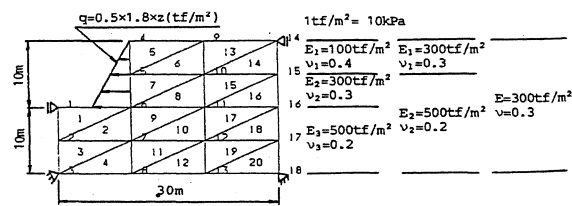
### 6.2 弾性定数の分布の推定

材料特性値の同定に関しては、構造の振動特性の同定に関連して多くの研究がなされている [122]。このほかにも、複合材料の変位の応答より弾性定数を同定する問題 [123]、弾性特性分布を推定する問題 [124] にカルマンフィルタが適用されている。拡張カルマンフィルタを用いて掘削地盤の弾性特性を推定した結果の例を図 20 に示す [123]。カルマンフィルタを適用した場合、解の収束の過程は、比較的安定している。

離散系の剛性関係式をもとに、材料特性値を推定する二つの方法が提案されている [125]。これらの方法を有限要素近似と組み合わせて、連続体の材料特性値分布の推定が行われている。ヤング率およびポアソン比の分布の推定にこれらの方法が適用されている [126]。

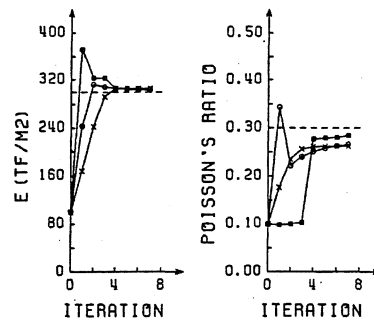
非均質材の評価は盛んで、インデンテーション法による薄膜弾性定数分布の推定、弾性波による層状材料の非破壊評価、超音波音響信号による膜構造材料の非破壊評価などの検討がある [127]–[129]。





(a) 解析対象

■—■ 荒井らの方法    ○—○ Kalman フィルター有限要素法  
×—× Cividini らの方法



(b) 推定の過程

図 20: 弾性定数分布の推定に対するカルマンフィルタの適用

材料の構成関係の同定にも逆問題解析が適用されており、粘弾性材料の緩和スペクトルなどの同定<sup>[130]</sup>、粘塑性材料の構成式のパラメータ同定<sup>[131]</sup>、複合材料の弾性定数同定<sup>[132]</sup>がなされている。

## 7 支配方程式逆問題とその取扱いの例

支配方程式を推定する逆問題として、現象に観測結果をもとに、これを合理的に説明する線形微分方程式を求める方法が提案されている<sup>[133]–[135]</sup>。これは、支配方程式の同定が係数の同定に帰着させるものである。すなわち、現象が常微分方程式で表されるとしたときのその階数を  $n$  として、非同次項をもたない常微分方程式は、

$$\frac{d^n \phi}{dx^n} + C_{n-1} \frac{d^{n-1} \phi}{dx^{n-1}} + \cdots + C_0 \phi = 0 \quad (9)$$

のように書ける。観測結果をもとに、微係数が評価できれば、未知係数  $C_i$  に関する関係式が得られる。このように支配微分方程式を推定する問題が、種々の観測結果をもとに得られた微係数より、微分方程式の階数と係数  $C_i$  の推定問題に置き換えられている。

この推定では、誤差を含む観測値からの微係数の数値的評価が問題となる。誤差の影響を軽減するためモリフィケーション法が適用されている。また、非線形方程式の推定や次式のような偏微分方程式の推定も行われている。

$$\frac{\partial^n \phi}{\partial x^n} + C_{n-1} \frac{\partial^n \phi}{\partial x^{n-1} \partial y} + \cdots + C_{0n} \frac{\partial^n \phi}{\partial y^n} + \cdots + C_{00} \phi = 0 \quad (10)$$

微分方程式の階数の推定には、情報量基準 AIC が用いられている。

## 8 おわりに

以上述べてきたように、逆問題には種々のものがある。どのような逆問題を取り上げることが重要である。次に、逆問題をどう設定するか問題にうまく解けるか否かがかかっており、ここに研究者の力量が発揮できる余地がある。逆解析に用いる手法にも、この問題ごとに先験情報の利用など工夫の余地が大きい。このような重要でありながら、工夫の余地の大きな逆問題に意欲的に取り組む研究者が増えることを祈念する。

## 参考文献

- [1] A.N. Tikhonov and V.Y. Arsenin, "Solutions of Ill-Posed Problems", (1977), John Wiley & Sons.
- [2] C.W. Groetsch, "The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind", (1984), Pitman.
- [3] G.M.L. Gladwell, "Inverse Problems in Vibration", (1986), Martinus Nijhoff Pub.

- [4] M.M. Lavrent'ev, V.G. Romonov and S.P. Shishat · skii, "Ill-Posed Problems of Mathematical Physics and Analysis", (1986), Amer. Math. Soc.
- [5] V.G. Romanov, "Inverse Problems of Mathematical Physics", (1987), VNU Sci. Press.
- [6] 久保司郎, 大路清嗣, "境界要素法の応用", 181-198 (1987), コロナ社.
- [7] 久保司郎, 機械の研究, **39**, 1071-0177(1987).
- [8] S. Kubo, JSME Int.J., Ser.I, **31**, 157-166(1988).
- [9] 日本機械学会シンポジウム「逆問題のコンピュータ手法とその応用」講演論文集, No.890-34(1989).
- [10] M. Bonnet and H.D. Bui, 境界要素法論文集, **6**, 179-197(1989).
- [11] 田中正隆, 日本機械学会第 67 期通常総会講演会資料集, Vol.D, 276-279(1990).
- [12] 瀬口靖幸, [11] の pp.285-288.
- [13] 山川 宏, [11] の pp.289-293.
- [14] 久保司郎, システム／制御／情報, **35**, 634-642(1991).
- [15] 日本機械学会 (編), "逆問題のコンピュータアナリシス", (1991), コロナ社.
- [16] M. Yamaguti, et al. (ed.), "Inverse Problems in Engineering Sciences", (1991), Springer.
- [17] 久保司郎, "逆問題", (1992), 培風館.
- [18] 武者利光, 岡本良夫, "逆問題とその解き方", (1992), オーム社.
- [19] S. Kubo (ed.), "Inverse Problems", (1993), Atlanta Technology Pub. Springer.
- [20] M. Tanaka and H.D. Bui (ed.), "Inverse Problems in Engineering Mechanics", 529(1993), Springer.
- [21] N. Zabarar, K.A. Woodbury and M. Raynaud (ed.), "Inverse Problems in Engineering: Theory and Practice", (1993), Engineering Foundation, ASME.
- [22] 数理科学, **403**, (1997).
- [23] 久保司郎, "逆に考え, 逆に解く", 日本機械学会 100 周年記念出版物 テクノライフ選書, (1997), オーム社.
- [24] 久保司郎, 応用数理, **7**, 181-195(1997).
- [25] D. Delaunay, Y. Yarny, and K.A. Woodbury (ed.), "Proc. 2nd Int. Conf. Inverse Problems in Engineering: Theory and Practice", (1998), Engineering Foundation.

- [26] M. Tanaka and G.S. Dulikravich (ed.), "Inverse Problems in Engineering Mechanics", (1998), Elsevier.
- [27] M. Tanaka and G.S. Dulikravich (ed.), "Inverse Problems in Engineering Mechanics II", (2000), Elsevier.
- [28] M. Tanaka and G.S. Dulikravich (ed.), "Inverse Problems in Engineering Mechanics III", (2001), Elsevier.
- [29] M. Tanaka (ed.), "Inverse Problems in Engineering Mechanics IV", (2003), Elsevier.
- [30] H.R.B. Orlande (ed.), "Inverse Problems in Engineering: Theory and Practice", (2002), Engineering Foundation.
- [31] 登坂宣好, 大西和榮, 山本昌宏, "逆問題の数理と解法", (1999), 東京大学出版会.
- [32] C.W. Groetsch (大西和榮, 田沼一実, 山本昌宏 訳), "はじめての逆問題", (1999), サイエンス社.
- [33] 久保司郎, 桑山真二郎, 大路清嗣, 日本機械学会論文集 (A 編), **61**-581, 169-176(1995).
- [34] T. Murai and Y. Kagawa, Int. J. Num. Methods in Eng., **23**, 35-47(1986).
- [35] 林 眞琴, 大高, 榎本, 佐々木, 菊地, 材料, **35**, 936-941(1986).
- [36] 矢川元基, 福田俊彦, 機械の研究, **36**, 27-32(1984).
- [37] 三好俊郎, 中野伸一, 日本機械学会論文集 (A 編), **52**-476, 1097-1102(1986).
- [38] L.F. Coffin, ASTM STP 969, 235-259(1988).
- [39] 大路清嗣, 久保司郎, 阪上隆英, 日本機械学会論文集 (A 編), **51**-467, 1818-1825(1985).
- [40] 久保司郎, 阪上隆英, 大路清嗣, 橋本, 松室, 日本機械学会論文集 (A 編), **54**498, 218-225(1988).
- [41] 阪上隆英, 久保司郎, 橋本 健, 山脇 弘, 大路清嗣, 日本機械学会論文集 (A 編), **53**-472, 1598-1605(1987).
- [42] 久保司郎, 阪上隆英, 大路清嗣, 橋本 健, 松室義隆, 日本機械学会論文集 (A 編), **54**-498, 218-225(1988).
- [43] 阪上隆英, 久保司郎, 大路清嗣, 山本, 中塚, 日本機械学会論文集 (A 編), **56**-521, 27-32(1990).
- [44] 大路清嗣, 久保司郎, 鹿兒島昌之, 今宿 匠, 日本機械学会論文集 (A 編), **57**-541, 2078-2084(1991).

- [45] 久保司郎, 阪上隆英, 塩沢大輝, 日本 AEM 学会誌, **11**, 73-81(2003).
- [46] 久保司郎, 阪上隆英, 大路清嗣, 日本機械学会論文集 (A 編), **55**-519, 2316-2319(1989).
- [47] S. Kubo, Lecture Notes in Numerical and Applied Analysis, **13**, 189-199(1994).
- [48] 大路清嗣, 久保司郎, 阪上隆英, 安田章二, 藤下健一, 材料, **39**, 1167-1172(1990).
- [49] S. Kubo, J. Sarasinpitak and K. Ohji, Mater. Sci. Research Int., **2**, 26-32(1996).
- [50] S. Kubo, T. Sakagami, H. Kitaoka and K. Ohji, Inverse Problems in Engineering Mechanics, (M. Tanaka, ed.), 91-100(1998), Elsevier.
- [51] S. Kubo, T. Sakagami, J. Sarasinpitak and K. Ohji, "Experimental Mechanics, Advances in Design, Testing & Analysis", I.M. Allison ed., Vol.2, Balkema, 933-938(1998).
- [52] S. Kubo, M. Kuchinishi, T. Sakagami, and S. Ioka, Int. J. Applied Electromagnetics and Mechanics, **15**, 261-267(2001/2002).
- [53] 坂 真澄, 大内, 阿部, 日本機械学会論文集 (A 編), **62**-593, 63-67(1996).
- [54] M. Saka, A. Oouchi and H. Abe, J. Pressure Vessels Tech., **118**, 198-202(1996).
- [55] Li S.-Q., S. Kubo, T. Sakagami, Liu Z.-X., Materials Sci. Research Int., **6**, 41-48(2000).
- [56] S. Shiozawa, S. Kubo, and T. Sakagami, Inverse Problems in Engineering: Theory and Practice, Engineering Foundation, ASME, **I**, 311-316(2002).
- [57] 塩沢大輝, 久保司郎, 阪上隆英, 奥野康輔, 材料, **51**, 926-931(2002).
- [58] 塩沢大輝, 久保司郎, 阪上隆英, 高木正彰, 日本機械学会講演論文集, No.03-26, 645-646(2003).
- [59] 塩沢大輝, 久保司郎, 阪上隆英, 岡崎俊介, 日本機械学会講演論文集, No.02-05, 497-498(2002).
- [60] 小島史男, 非破壊検査, **47**, 105-113(1998).
- [61] 黄 皓宇, 福富広幸, 高木敏行, 日本機械学会論文集 (A 編), **65**-638, 2024-2031(1999).
- [62] 小島史男, 日本 AEM 学会誌, **11**, 66-72(2003).
- [63] 田中正隆, 山際孝次, 日本機械学会論文集 (A 編), **54**-501, 1054-1059(1988).
- [64] 田中正隆, 中村, 中野, 日本機械学会論文集 (A 編), **56**-528, 1900-1905(1990).

- [65] 古口日出男, 渡部, 矢田, 日本機械学会第6回計算力学講演会講演論文集, No.930-71, 161-162(1993).
- [66] 中村正行, 田中, 望月, 日本機械学会第6回計算力学講演会講演論文集, No.930-71, 173-174(1993).
- [67] 小島史男, 非破壊検査, **45**, 249-253(1996).
- [68] 阪上隆英, 小倉敬二, D.D. Burlcigh, J.W.M. Spiccr, 非破壊検査, **45**, 788-797(1996).
- [69] 阪上隆英, 久保司郎, 非破壊検査, **48**, 653-660(1999).
- [70] 阪上隆英, 久保司郎, 三木康弘, 非破壊検査, **48**, 688-696(1999).
- [71] Y. Niwa and S. Hirose, Proc. Japan Soc. Civil Eng., Struct. Eng./Earthq. Eng., **3**, 267s-275s(1986).
- [72] S. Hirose, "Inverse Problems" (S. Kubo, ed.), 99-108(1993), Atlanta Technology Pub.
- [73] T. Fukui, "Inverse Problems in Engineering Mechanics" (M. Tanaka and H.D. Bui, ed.), 201-212(1993), Springer.
- [74] 北原道弘, 非破壊検査, **47**, 234-237(1998).
- [75] 上田光宏, 非破壊検査, **47**, 238-242(1998).
- [76] 山脇 寿, 非破壊検査, **47**, 243-248(1998).
- [77] 大石篤哉, 山田勝稔, 吉村 忍, 矢川元基, 日本機械学会論文集 (A編), **62**-598, 1479-1487(1996).
- [78] 大石篤哉, 山田勝稔, 吉村 忍, 矢川元基, 日本機械学会論文集 (A編), **62**-602, 2350-2357(1996).
- [79] 星川 洋, 非破壊検査, **45**, 848-853(1996).
- [80] K. Kishimoto, H. Miyasaka and S. Aoki, JSME Int. J., Ser.I, **32**, 256(1989).
- [81] 青木, 浦井, 材料, **40**, 619(1991).
- [82] 青木 繁, 天谷賢治, 富樫 潤, 日本機械学会論文集 (A編), **59**-562, 1562-1567(1993).
- [83] 久保司郎, 桑山真二郎, 大路清嗣, 日本機械学会論文集 (A編), **61**-581, 169-176(1995).
- [84] S. Kubo, T. Takahashi, and K. Ohji, Inverse Problems in Engineering Mechanics (M. Tanaka and G.S. Dulikravich ed.), 337(1998), Elsevier.

- [85] V.A. Kolov, V.G. Maz'ya, and A.V. Formin, *Comp. Maths. Math. Phys.*, **31**, 45-52(1991).
- [86] D. Lesnic, L.Elliott, and D.B. Ingham, *Eng. Analysis with Boundary Elements*, **20**, 123-133(1997).
- [87] S. Kubo, and A. Furukawa, *Inverse Problems in Engineering Mechanics II*, M. Tanaka and G.S. Dulikravich ed., Elsevier, 149-158(2000).
- [88] S. Kubo, and A. Furukawa, *Inverse Problems in Engineering Mechanics III*, M. Tanaka and G.S. Dulikravich ed., Elsevier, 77-86(2002).
- [89] 尾田, 品田, 日本機械学会論文集 (A編), **53**, 1614(1987).
- [90] 尾田十八, 本 悟, 日本機械学会論文集 (A編), **56**-526, 1479-1484(1990).
- [91] H. Koguchi, T. Tomishima and T. Yada, *Int. J. Pres. Ves. Piping*, **44**, 49(1990).
- [92] S. Kubo, and K. Ohji, *Computational Methods for Solution of Inverse Problems in Mechanics*, ASME, AMD-Vol.228, (L.G. Olson and S. Saigal ed.), 41(1998).
- [93] D.S. Schnur and N. Zabaras, *Int. J. Numerical Methods in Eng.*, **30**, 57(1990).
- [94] Z. Gao and T. Mura, *Int. J. Solids Struct.*, **25**, 901(1989).
- [95] S. Kubo, and M. Kawaguchi, *Inverse Problems in Engineering Mechanics* (M. Tanaka and G.S. Dulikravich ed.), 353(1998), Elsevier.
- [96] M. Tanaka, M. Nakamura, and R. Ochiai, "Inverse Problems in Engineering Mechanics" (M. Tanaka and H.D. Bui, ed.), 383(1993), Springer.
- [97] W. Waldman, T.G. Ryall and R. Jones, *Computers & Struct.*, **36**, 553(1990).
- [98] 村上敬宜, 吉村正昭, 日本機械学会論文集 (A編), **61**, 2482(1995).
- [99] 岸本喜久雄, 井上裕嗣, 新保英男, 渋谷壽一, 日本機械学会論文集 (A編), **62**-597, 1204-1211(1996).
- [100] 早房敬祐, 岸本喜久雄, 井上裕嗣, 渋谷壽一, 日本機械学会講演論文集, No.96-25, 109(1996).
- [101] S.A. Lukasiewicz, M. Stanuszek and J.A. Czyz, *Exp. Mech.*, **33**, 139(1993).
- [102] M.A. Sutton, et al., *Exp. Mech.*, **36**, 99(1996).
- [103] 田中 博, *医用電子と生体工学*, **23**, 147-158(1985).
- [104] 村井忠邦, 加川幸雄, *電子通信学会論文誌*, **J64C**, 223-230(1981).
- [105] 上田幸雄, 福田, 谷川, 日本造船学会論文集, **145**, 203-211(1979).

- [106] J. Kihara, G. Shen, T. Yamauchi, H. Mimura, H. Makino and B. Liu, "Boundary Elements", Proc. 5th Int. Conf., Hiroshima, 393-405(1983), Springer.
- [107] 馬 曉鵬, 木原諄二, 第 1 回境界要素法シンポジウム発表論文集, 97-102(1984).
- [108] 上田幸雄, 金, 梅国, 溶接学会論文集, **3**, 611-616(1985).
- [109] E.F. Rybicki and J.R. Shadley, Trans. ASME, J. Eng. Mater. Tech., **108**, 99-106(1986).
- [110] 富島俊彦, 矢田敏夫, 日本機械学会論文集 (A 編), **54**-500, 642-647(1988).
- [111] Z. Gao and T. Mura, Trans. ASME, J. Appl. Mech., **56**, 508-513(1989).
- [112] H. Koguchi, T. Tomishima and T. Yada, Int. J. Pres. Ves. Piping, **44**, 49-66(1990).
- [113] 久保司郎, 辻 昌宏, 大路清嗣, 日本機械学会論文集 (A 編), **54**-501, 892-899(1988).
- [114] K. Ohji, S. Kubo and M. Tsuji, "Fatigue 90", 4th Int. Conf. on Fatigue and Fatigue Threshold, Vol.IV, 2377-2382(1990).
- [115] S. Kubo, S. Hiramatsu, M. Tsuji and K. Ohji, "Proc. 2nd Int. Conf. Inverse Problems in Engineering: Theory and Practice" (D. Delaunay, ed.), Engineering Foundation, ASME, 237-244(1998).
- [116] 苦米地重尚, 久保司郎, 辻 昌宏, 山本成人, 日本機械学会講演論文集, No.96-10, Vol.A, 533-534(1996).
- [117] T. Kishi, Z. Metallkunde, **76**-7, 512-519(1985).
- [118] 久保司郎, 大中幸三郎, 大路清嗣, 日本機械学会論文集 (A 編), **54**-503, 1329-1334(1988).
- [119] 久保司郎, 大中幸三郎, 大路清嗣, 日本機械学会論文集 (A 編), **60**-575, 1638-1643(1994).
- [120] S. Kubo and K. Ohji, Inverse Problems in Engineering Mechanics, H.D. Bui et al. ed., Balkema, 289-293(1994).
- [121] 上田 整, 日本機械学会論文集 (A 編), **54**-500, 638-641(1988).
- [122] A. Nagamatsu, JSME Int. J., **30**, 1526(1987).
- [123] 村上 章, 長谷川高士, 土木学会論文集, **388-III-8**, 227-235(1987).
- [124] W.M.G. Courage, P.J.G. Schreurs and J.D. Janssen, Computers and Struct., **34**, 231(1990).
- [125] 久保, 大路, 日本機械学会論文集 (A 編), **57**, 2257(1991).



- [126] S. Kubo, K. Ohji and K. Konishi, "Inverse Problems" (S. Kubo, ed.), 212(1993), Atlanta Technology Pub.
- [127] 古口日出男, 南雲悦男, 日本機械学会論文集 (A編), **62**, 513(1996).
- [128] 井原郁夫, 古口日出男, 相澤龍彦, 木原諄二, 日本機械学会講演論文集, No.96-10, Vol.A, 535(1996).
- [129] Tomas Novotny, 松本英治, 柴田俊忍, 日本機械学会講演論文集, No.96-10, Vol.A, 529(1996).
- [130] Yu.G. Yanovsky, et al., Int. J. Eng. Sci., **34**, 1221(1996).
- [131] O. Ghouati and J.C. Gelin, Inverse Problems in Engineering: Theory and Practice (D. Delaunay, Y. Jarny, and K.A. Woodbury ed.), 115(1998), Engineering Foundation, ASME.
- [132] Zhiqing Zou and Michael Link, Inverse Problems in Eng., **3**, 145(1996).
- [133] 久保司郎, 大路清嗣, 塩尻明夫, 日本機械学会論文集 (A編), **57**, 2250(1991).
- [134] Shiro Kubo, Kiyotsugu Ohji, and Akio Shiojiri, Inverse Problems in Engineering: Theory and Practice, Eds. N. Zabaras, K.A. Woodbury and M. Raynaud, ASME, 283-289(1993).
- [135] Shiro Kubo, Koji Fukumoto, and Kiyotsugu Ohji, Inverse Problems in Engineering: Theory and Practice, Eds. D. Delaunay, Y. Jarny and K.A. Woodbury, ASME, 245-252(1998).

# 逆問題の解の直接再構成

池 畠 優

群馬大学工学部共通講座

376-8515 群馬県桐生市天神町一丁目 5 番 1 号

2005 年 9 月 6 日

## 目次

### 序

### 第一章 体積（面積）のある不連続性を抽出すること

#### 1 境界値逆問題の原型

##### 1.1. 支配方程式の導出

##### 1.2. Dirichlet-to-Neumann 写像

#### 2 囲い込み法と不連続性の凸包の抽出

##### 2.1. 囲い込み法の始まり

##### 2.2. 単純化された囲い込み法

#### 3 探針法と不連続性そのものの抽出

##### 3.1. 針と針列

##### 3.2. 探針法の A 面

##### 3.3. 探針法の B 面

#### 4 囲い込み法と Mittag-Leffler の関数

##### 4.1. 囲い込み法の一般化

##### 4.2. Mittag-Leffler の関数の漸近挙動

##### 4.3. 不連続性の可視部の抽出

#### 5 因数分解法

##### 5.1. 因数分解公式

##### 5.2. 因数分解公式の一つの帰結

### 第二章 体積（面積）のない不連続性を抽出すること

#### 6 探針法と亀裂の逆問題

##### 6.1. 亀裂の逆問題

##### 6.2. 指示列および指示関数

##### 6.3. 反射解の爆発

#### 7 囲い込み法と逆散乱問題

##### 7.1. 体積のない障害物による音波の逆散乱問題

##### 7.2. 散乱振幅から凸包を抽出すること

##### 7.3. Herglotz の波動関数、Vekua 変換および指示関数の漸近展開

##### 7.4. 視野が限定されたときの先験情報の役割

### 文献

## 序

逆問題の研究者の**最終目標**は何か。それは観測データから未知のものの情報を抽出すること（これを単に逆問題の解の再構成ということにする）である。逆問題を研究するのはこれを片時も忘れてはならない。言い換えれば、今自分がやっていることがその最終目標にどのようにかわるのかを常に意識し自問自答することが必要とされる。

逆問題の数学的研究はいくつかの段階に分かれるという。**第一段階**は、観測データと未知のものの数学モデルの設定。観測データは物理量で与えられる故、多くのモデルは**支配方程式**と呼ばれる偏微分方程式で記述され、観測データはその解を使って、未知のものはその偏微分方程式の境界条件、係数、解のある場所での値などとしてモデルに取り込まれる。このときの観測データを**理想化された観測データ**と呼ぶことにする。

**第二段階**は、通常は、この数学モデルの中で、理想化された観測データが未知のものを一意的に決定するかという**逆問題の解の一意性**の問題、さらには理想化された観測データが互いに何らかの意味で近ければ、それを産み出すもととなった未知のもの同士も何らかの意味で近いのかという**逆問題の解の安定性**の問題に取り組むことである。これらの問題の理解および解は、適切な評価関数と正則化パラメタの選択に知見を与え、それを使った、最終目標である逆問題の解の構成につながるというのがその研究の意義である。解は、解の候補者からなる適当なクラスで、選択された評価関数を最適化するものとして与えられる。この研究の**進め方**は、普遍性があり、個々の逆問題の個性に左右されないという利点がある。もちろん、一意性と安定性の問題の解決自体には個々の逆問題の個性の把握が必要であることは言うまでもない。この進め方は**逆問題の解の間接再構成**呼んでもよいと思われる。

一方、この講義では、上述の第二段階に、逆問題の解の情報を理想化された観測データから直接解析的に抽出する公式（**解の情報の抽出公式**）の発見をもって来る。これはその公式にあらわれる解の情報が理想化された観測データからどのように決定されるのかという**からくり**を明らかにする。例をあげるまでもなく、このような公式を得ることはおよそ**逆問題の研究者の夢**ではないだろうか。例え一意性がなくても、解の情報の一部でも抽出できれば有用であることは明らかであろう。ただし、これで終わりではない。実際の観測データは誤差やノイズが混入しており、それをそのまま公式に代入しても、途中で計算が破綻する。これが逆問題の特徴の一つの表現であり、逆問題の解の間接再構成においても同じ問題が発生し、いつ計算を止めるかが重要になってくる。これは公式の正則化の問題であり、逆問題の解の間接再構成における評価関数の正則化に対応している。この講義ではこの進め方を**逆問題の解の直接再構成**と呼ぶことにする。ただこの進め方は逆問題の解の情報の抽出公式の**発見**がなければなにも始まらないことはいうまでもない。しかし、これは、個々の逆問題の個性を、もっと深く把握しようという方向の研究も促す。個々の逆問題を、あたかも個々の特殊関数と思い、その個性を追求する。その結晶化したものが、解の情報の抽出公式ではないだろうか。

このほぼ十年間、逆問題の研究とくに、亀裂、障害物、介在物、空洞などの媒質中の**不連続性**の抽出問題に関する研究が大きく変貌した。従来はこの解の間接再構成の研究がほとんどであった。その代表的な文献として、Alessandrini-DiBenedetto[2]、Isakov[52]、Friedman-Isakov[52]、Friedman-Vogelius[20]らの論文があげられるであろう。しかし今や解の直接再構成の研究が雨後の竹の子のようにあらわれて群雄割拠(?)の戦国時代の様相を呈し、まさに室町幕府は崩壊したのである。著者は1997年4月**仮想的な針**（探針）をつかって不連続性を抽出するアイデアを発見しそれを**探針法**と名づけた。これをまとめた論文は、紆余曲折を経て翌年偏微分方程式に関する専門誌 Comm. PDE. に掲載され

た ([24])。またこのアイデアを障害物による音波の散乱の逆問題へ応用し、入射波として平面波を与える場合を扱った論文がその年逆問題に関する専門誌 *Inverse Problems* に掲載され ([25])、少し遅れて、点源を発する球面波を入射波とした場合を扱った論文が波動現象の専門誌 *WAVE MOTION* に掲載された ([26])。これは凸性などの条件を仮定することなしに障害物そのものの**再構成公式**を与える世界で最初のものであり、著者のこの問題における研究状況に対する長年の知的欲求不満の解消になった。しかし、批判もあった。それは探針法の基礎が **Runge の近似定理**であって、構成的ではないのではという指摘である。批判者が言う**構成的であることの定義**を私は知らないが、1998 年、不連続性の凸包を **Runge の近似定理**を使用しないで抽出する方法として、著者は**囲い込み法**を提唱した ([29])。この囲い込み法の考え方は、Sylvester-Uhlmann[69] によって構成された、**定常 Schrödinger 方程式の複素幾何光学解** (Complex geometrical optics solution) の、新しい応用を与え、それは**逆源泉問題** [28] や定常 Schrödinger 方程式に対する **Cauchy 問題** [34, 51] へも応用された。

2000 年頃までのこの二つの方法の研究状況については、東京都立大学での集中講義をまとめたノート [30] を見られたい。不思議なことに、時期的に非常に近接して、他のいくつかの基本的アイデアも提唱された。障害物による音波などの散乱問題において提唱された Kirsch の**因数分解法 (Factorization method)** [54] あるいは Potthast の**特異源泉法 (Singular sources method)** [66] などがそれである。現在特異源泉法は探針法の、この講義で説明するところの A 面と本質的には同じであることがわかっている。なお初期に提唱された Colton-Kirsch の **Linear sampling method** [11] はその数学的正当化に問題があったが、最近 Tilo[71] により満足のゆく形に正当化された。

この講義はノート [30] の続編に位置づけられるであろう。そしてこの講義の構成であるが、大きく二つに分けてある。

第一章では、空洞のような体積をもった不連続性を観測データから抽出する問題をとりあげ、囲い込み法、探針法および因数分解法について、解説した。支配方程式は **Laplace 方程式**であり、いきなり問題に入っていくてかつアイデアの核心が掴みやすいということで、この問題を選択したが、最低限必要な順問題についての知識については、簡単にまとめておいた。トレース定理についてはそれが正しく使えればよしとし、必要があれば例えば Grisvard[21] の本を参照すればよい。一節および二節の内容のいくつかの部分はノート [30] にもふくまれているが、支配方程式の導出についての簡単な説明そして単純化された囲い込み法 ([27]) のキーとなる補題の幾分くどい証明はそのノートにはふくまれていない。また最近の成果 [39] で得られた知見として、**スペクトルの逆問題**への応用についての注意も与えておいた。三節の探針法は、最近出版した探針法の見直しに関する論文 [43] で得られた新しい知見にもとづいて書かれていて、[30] とは異なる。[24] の探針法は、この講義によれば**二つの側面の一つ**しか見ていなかったことがわかるであろう。四節では、**囲い込み法の一般化**についての成果 ([40]) を紹介する。**Mittag-Leffler の関数**の漸近挙動についての証明と不連続性の抽出公式の証明を詳細に記述した。前者は Bateman の有名な本 [3] を参考にしたことを注意しておこう。著者は数年前までこの特殊関数の名前すら知らなかったのであるが、東大の図書館でウズベキスタンの Yarmukhamedov の Laplace 方程式に対する Cauchy 問題についての、およそ現代の数学とは全く対極的な異色の論文 [75] を偶然見つけ、その存在と性質を知ることとなった。なお彼の興味深い研究結果は、[33] で、囲い込み法と組み合わせられて、介在物の抽出問題へ応用された。Mittag-Leffler の関数はおよそ百年前に Mittag-Leffler により導入され、指数関数のある意味での一般化で、実は確率論や積分方程式などへの応用もあることを後で知ったが、このような問題に応用できたことは、著者にとって、大変な喜びであった。

さて、この講義の目的は先端的話題を紹介することであった。それゆえ、探針法や囲い込

み法と対照的な、Kirschによって導入された**因数分解法**についての解説を含まなければ完全ではないであろう。五節では、彼の方法を空洞の問題に適用し、その中心である**因数分解公式**の導出方法とその使い方に絞って解説した。

第二章は、探針法および囲い込み法を、亀裂のような体積がない不連続性を観測データから抽出する問題を取りあげた。六節では亀裂の逆問題へ探針法を応用した。亀裂は体積がないぶん、そのままでは困難が増すが、**うまい試験関数の選択**により切り抜けられることを示した。ここはプレプリント [45] にもとづいていることを注意しておこう。七節では、囲い込み法を、体積のない障害物による、音波の散乱の逆問題へ応用した。支配方程式が Laplace から Helmholtz にかわっただけと思われるかもしれないが、そのままでは解析は格段に複雑になることがわかる。ところが**パラメタのうまい変更**により、結果としてすべてがうまくいった。この辺はまさに、初等幾何学における**うまい補助線の発見**の喜びに通じるものであった。さらに古典的な Vekua の変換公式、Herglotz の波動関数などがからみ、この種の問題における基本的な部分の紹介にもなったのではないかと思っているが、いかがであろうか。ただこの節では肝心の部分の証明は省略しているので、興味のある方は原論文 [39] およびプレプリント [41] にあたっていただきたい。

いくつかの節の終わりには、著者が、常日頃、その解答を知りたい問題について、そのうちのいくつかを注意しておいた。この機会を通じてそれらについて何かみなさんから知見が得られれば幸いである。探針法および囲い込み法は、すでに、共同研究者、他のグループにより数値実験がなされ ([6, 10, 18, 47, 49, 50])、他の問題へ応用され ([53, 63, 64])、第三者による Survey paper [67] も存在する。さらに、この考え方を基礎にした論文もあらわれている ([65])。このように、**探針法と囲い込み法**の考え方は、受け入れられつつあるようにみえる。

2005年6月30日 桐生にて

## 第一章 体積（面積）のある不連続性を抽出すること

この章では与えられた物体内に発生した空洞のような体積（面積）のある欠陥の位置や形についての情報を観測データから抽出する公式とその背後にある考え方を具体的な境界値逆問題を通して紹介する。

### 1 境界値逆問題の原型

この節では、電荷が動くことのできる物体すなわち**導体**の内部に何等かの理由で発生した**空洞**の位置および形状を物体表面を通して与えられた定常電流密度と対応する表面電位から抽出する問題を **Dirichlet-to-Neumann 写像**を使って定式化する。

#### 1.1. 支配方程式の導出

出発点は Maxwell の方程式系である。空洞がない場合の導体を  $\Omega$  であらわす。 $\rho$ 、 $\epsilon$ 、 $\mu$  および  $J$  でそれぞれ、その導体まで込めた空間の電荷密度 (charge density)、誘電率 (electric permittivity)、透磁率 (magnetic permeability) および電流密度 (current) をあらわす。電磁気学によればこのとき発生した電場  $E$  (electric field) および磁場  $H$  (magnetic field) は次の Maxwell の方程式系を満たす。

$$\nabla \cdot (\epsilon E) = \rho, \quad (1.1)$$

$$\nabla \times E = -\frac{\partial}{\partial t}(\mu H), \quad (1.2)$$

$$\nabla \cdot (\mu H) = 0, \quad (1.3)$$

$$\nabla \times H = J + \frac{\partial}{\partial t}(\epsilon E). \quad (1.4)$$

(1.1) は Gauss の法則、(1.2) は Faraday の法則、(1.3) は磁気単極子の非存在、(1.4) は Ampere の法則である。(1.4) と (1.1) より電荷保存の法則

$$\nabla \cdot J + \frac{\partial}{\partial t}\rho = 0 \quad (1.5)$$

を得る。いま電場、磁場、電流密度および電荷密度すべて時間に関して一定であるとする。このとき (1.5) から

$$\nabla \cdot J = 0 \quad (1.6)$$

を得る。これが Kirchhoff の法則である。一方 (1.2) より

$$\nabla \times E = 0$$

を得る。このとき

$$E = -\nabla u \quad (1.7)$$

を満たす  $u$  が存在する。ここで導体内部で Ohm の法則

$$J = \sigma E \quad (1.8)$$

が成り立っていると仮定する (いいかえればこの法則が成り立つ導体を考えている)。 $\sigma$  はこの導体の導電率 (conductivity) と呼ばれる。(1.6)、(1.7) および (1.8) より導体内部における導電場の方程式

$$\nabla \cdot \sigma \nabla u = 0 \text{ in } \Omega \quad (1.9)$$

を得る。このとき導体表面における電流密度は

$$j = -\sigma \nabla u \cdot \nu|_{\partial\Omega}$$

で与えられる。ここで  $\nu$  は  $\partial\Omega$  に対する外向き単位法線ベクトル場である。 $u$  は電位 (voltage potential) と呼ばれ、導体表面における電位は

$$f = u|_{\partial\Omega}$$

で与えられる。 $\sigma$  に対する正值性の仮定のもとで、導体表面における電位は導体表面における電流密度を一意的に決定する。この講義では、対応  $f \mapsto j$  を導電率  $\sigma$  の導体  $\Omega$  に対する Dirichlet-to-Neumann 写像と呼び  $\Lambda_\sigma$  であらわす。今から 4 半世紀前数学者 A.P. Calderón は、 $\Lambda_\sigma$  によって  $\sigma$  が一意的に決定されるかそして決定されるときどのように  $\sigma$  を  $\Lambda_\sigma$  を使って計算するかという問題を定式化した ([8])。この Calderón の問題は、その導出からわかるように、導体表面における電位と電流密度分布から導体内部の導電率を知ろうとする電気インピーダンストモグラフィの連続体モデルの基礎にかかわり、その後の展開が示すように ([73])、数学としてみても極めて興味深い問題である。

さてこの第一章では既知の導電率  $\sigma$  を持った導体内部に、なんらかの理由で発生した空洞  $D$  について、その位置と形についての情報を導体表面における電位と電流密度分布

から知ろうとする問題を取り上げる。 $\sigma$  は空洞発生前の導体の導電率である。空洞の表面  $\partial D$  を通して電荷の出入りが無いということを仮定すると  $\partial D$  における境界条件

$$-\sigma \nabla u \cdot \nu = 0 \text{ on } \partial D$$

を得る。このとき (1.9) から

$$\nabla \cdot \sigma \nabla u = 0 \text{ in } \Omega \setminus \overline{D} \quad (1.10)$$

を得る。これは形式的には Calderón の問題において、 $D$  上  $\sigma = 0$  とすることに対応する。これは  $D$  は極端に電流が流れにくい領域であることを意味する。

$\sigma$  に対する  $\Omega \setminus \overline{D}$  における正值性の仮定のもとで、この場合も導体表面における電位は導体表面における電流密度を一意的に決定する。このとき対応  $f \mapsto j$  を Dirichlet-to-Neumann 写像と呼び  $\Lambda_D$  であらわす。

この章を通して我々は、次の問題を追及する。

• 導電率  $\sigma$  は一定の定数で既知であるとする。 $\Lambda_D$  から  $D$  についての位置あるいは形についての情報を抽出すること。

このとき、(1.10) は Laplace 方程式

$$\Delta u = 0 \text{ in } \Omega \setminus \overline{D}$$

となる。この講義では、この問題の 2 次元版を考える。同じ記号を使用することにする。この場合  $\Omega$  は導体の注目している断面となり、電流密度のこの断面に垂直方向の成分は無視できるとして対応する 2 次元の問題が得られる。

## 1.2. Dirichlet-to-Neumann 写像

さてこれから先は数学の出番である。 $\Omega$  は単に漠然と導体の注目している断面としたが、数学の言葉ではっきりとその居場所を規定しよう。この章を通して、 $\Omega$  はその境界が Lipschitz である  $\mathbf{R}^2$  の中の有界な領域（連結開集合）であるとする。その中に含まれる未知の空洞の断面  $D$  は  $\Omega$  の開集合でその閉包  $\overline{D}$  が  $\Omega$  に含まれ、 $\Omega \setminus \overline{D}$  が連結であると仮定する。そしてさらに、 $D$  は、その閉包が互いに素である、Lipschitz 境界をもつ有界な領域の有限個の和集合

$$D = D_1 \cup \dots \cup D_N$$

で与えられていると仮定する。 $\nu$  で  $\Omega \setminus \overline{D}$  に対する  $\partial(\Omega \setminus \overline{D})$  上の外向き単位法線ベクトル場をあらわす。

ここでは、Laplace 方程式に対する二つの境界値問題の弱解の定義を与え、その存在と一意性についての知見を述べる。

**定義 1.1.** 与えられた  $f \in H^{1/2}(\partial\Omega)$  に対し  $u \in H^1(\Omega \setminus \overline{D})$  が次の境界値問題

$$\begin{aligned} \Delta u &= 0 \text{ in } \Omega \setminus \overline{D}, \\ \frac{\partial u}{\partial \nu} &= 0 \text{ on } \partial D, \\ u &= f \text{ on } \partial\Omega \end{aligned} \quad (1.11)$$

の弱解であるとは、 $u$  の  $\partial\Omega$  上へのトレースが  $f$  と一致しかつその  $\partial\Omega$  上へのトレースが消えるすべての  $\varphi \in H^1(\Omega \setminus \overline{D})$  に対して

$$\int_{\Omega \setminus \overline{D}} \nabla u \cdot \nabla \varphi dx = 0$$

が成り立つときをいう。また  $v \in H^1(\Omega)$  が次の境界値問題

$$\begin{aligned}\Delta v &= 0 \text{ in } \Omega, \\ v &= f \text{ on } \partial\Omega\end{aligned}\tag{1.12}$$

の弱解であるとは、 $v$  の  $\partial\Omega$  上へのトレースが  $f$  と一致しかつその  $\partial\Omega$  上へのトレースが消えるすべての  $\varphi \in H^1(\Omega)$  に対して

$$\int_{\Omega} \nabla v \cdot \nabla \varphi dx = 0$$

が成り立つときをいう。

境界値問題 (1.11)、(1.12) それぞれの弱解の存在と一意性は基本的であり以下のよ  
うにまとめられる。

**命題 1.1.** 与えられた  $f \in H^{1/2}(\partial\Omega)$  に対し (1.12) の弱解  $v \in H^1(\Omega)$  が唯一つ存在する。  
さらに、 $f$  には独立な正の定数  $C_1 = C_1(\Omega)$  が存在して評価

$$\|v\|_{H^1(\Omega)} \leq C_1 \|f\|_{H^{1/2}(\partial\Omega)}\tag{1.13}$$

が成り立つ。

**命題 1.2.** 与えられた  $f \in H^{1/2}(\partial\Omega)$  に対し (1.11) の弱解  $u \in H^1(\Omega \setminus \overline{D})$  が唯一つ存在す  
る。さらに、 $f$  には独立な正の定数  $C_2 = C(D)$  が存在して評価

$$\|u - v\|_{H^1(\Omega \setminus \overline{D})} \leq C_2 \|\nabla v\|_{L^2(D)}\tag{1.14}$$

が成り立つ。ここに  $v \in H^1(\Omega)$  は  $f$  に対する (1.12) の弱解である。

さて Dirichlet-to-Neumann 写像を定義し、この章を通して考察される逆問題を定式化  
しよう。

**定義 1.2.**  $f \in H^{1/2}(\partial\Omega)$  を任意に与える。

(1)  $H^{1/2}(\partial\Omega)$  上の有界線形汎関数  $\Lambda_0 f$  を公式

$$\langle \Lambda_0 f, g \rangle = \int_{\Omega} \nabla v \cdot \nabla \eta dx, \quad g \in H^{1/2}(\partial\Omega)$$

により定義する。ここで  $v \in H^1(\Omega)$  は (1.12) の弱解および  $\eta \in H^1(\Omega)$  はその  $\partial\Omega$  上へのト  
レースが  $g$  である任意の関数である。写像  $f \mapsto \Lambda_0 f$  を背景になる Dirichlet-to-Neumann  
写像と呼ぶことにする。

(2)  $H^{1/2}(\partial\Omega)$  上の有界線形汎関数  $\Lambda_D f$  を公式

$$\langle \Lambda_D f, g \rangle = \int_{\Omega \setminus \overline{D}} \nabla u \cdot \nabla \varphi dx, \quad g \in H^{1/2}(\partial\Omega)$$

により定義する。ここで  $u \in H^1(\Omega)$  は (1.11) の弱解および  $\varphi \in H^1(\Omega \setminus \overline{D})$  はその  $\partial\Omega$  上へ  
のトレースが  $g$  である任意の関数である。写像  $f \mapsto \Lambda_D f$  を単に Dirichlet-to-Neumann  
写像と呼ぶことにする。



もし  $\partial\Omega$  と  $f$  の両方が滑らかであれば、楕円型方程式の解の境界までこめた正則性の結果より、 $u$  および  $v$  の両方とも  $\partial\Omega$  の  $\bar{\Omega}$  に関するある相対近傍で滑らかであり古典的表現

$$\Lambda_D f = \frac{\partial u}{\partial \nu} \Big|_{\partial\Omega};$$

$$\Lambda_0 f = \frac{\partial v}{\partial \nu} \Big|_{\partial\Omega}$$

を得る。

この章ではつぎの問題を考察する。

**逆問題 1.1.**  $\Lambda_D$  あるいはその一部から  $D$  の位置および形についての情報を抽出せよ。

## 2 囲い込み法と不連続性の凸包の抽出

2次元のすべての単位ベクトルからなる集合を  $S^1$  であらわす。これは原点を中心とする単位円周と同一視される。与えられた  $\omega = (\omega_1, \omega_2) \in S^1$  に対し  $\omega$  に直交する単位ベクトルは二つある。そのうち  $\omega^\perp = (\omega_2, -\omega_1) \in S^1$  とおく。

この節では、おおきなパラメタ  $\tau(>0)$  を持つ調和関数

$$v_\tau(x; \omega) = e^{\tau x \cdot (\omega + i\omega^\perp)}$$

が中心的役割を果たす。

### 2.1. 囲い込み法の始まり

**定義 2.1.** 指示関数  $I_\omega(\tau; t)$  を公式

$$I_\omega(\tau; t) = \langle (\Lambda_0 - \Lambda_D)(e^{-\tau t} v_\tau(\cdot; \omega)|_{\partial\Omega}), \overline{e^{-\tau t} v_\tau(\cdot; \omega)} \rangle$$

により定義する。

この部分節では、上で定義した指示関数の  $\tau \rightarrow \infty$  のときの漸近挙動から  $D$  の凸包が抽出されることを示す。

$D$  の凸包はその支持関数

$$h_D(\omega) = \sup_{x \in D} x \cdot \omega, \quad \omega \in S^1$$

から次のようにもとまる：

$$\bigcap_{\omega \in S^1} \{x \in \mathbf{R}^2 \mid x \cdot \omega < h_D(\omega)\}.$$

それゆえ、 $D$  の支持関数が指示関数から計算できることを示せばよい。

ここで極めて有用な不等式系を述べておく。その不等式系は、背景になる Dirichlet-to-Neumann 写像  $\Lambda_0$  と Dirichlet-to-Neumann 写像  $\Lambda_D$  のずれを表現していて基礎的である。

**命題 2.1.** 与えられた  $f$  に対し  $v \in H^1(\Omega)$  を (1.2) の弱解とする。このとき  $D$  のみに依存する正の定数  $C_D$  が存在し不等式系

$$\int_D |\nabla v|^2 dx \leq \langle (\Lambda_0 - \Lambda_D)f, \bar{f} \rangle \leq C_D \int_D |\nabla v|^2 dx \quad (2.1)$$

が成り立つ。

証明  $u$  を  $f$  に対する (1.11) の弱解とする。Dirichlet-to-Neumann 写像の定義を注意深く使くと等式

$$\langle (\Lambda_0 - \Lambda_D)f, \bar{f} \rangle = \int_{\Omega \setminus \bar{D}} |\nabla(u - v)|^2 dx + \int_D |\nabla v|^2 dx \quad (2.2)$$

を得る。ここは演習問題としよう。このとき (2.1) は (1.14) からの直接の帰結である。

□

囲い込み法は次の定理で始まった。

定理 2.2. 公式

$$\lim_{\tau \rightarrow \infty} \frac{\log |I_\omega(\tau; 0)|}{2\tau} = h_D(\omega), \quad (2.3)$$

が成り立つ。さらに以下の主張は正しい：

- もし  $t > h_D(\omega)$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} I_\omega(\tau; t) = 0;$$

- もし  $t < h_D(\omega)$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} I_\omega(\tau; t) = \infty;$$

- もし  $t = h_D(\omega)$  ならば、そのとき

$$\liminf_{\tau \rightarrow \infty} I_\omega(\tau; t) > 0.$$

証明 自明な等式

$$I_\omega(\tau; t) = e^{2\tau(h_D(\omega) - t)} I_\omega(\tau; h_D(\omega)) \quad \forall t \in \mathbf{R}$$

によればすべては  $t = h_D(\omega)$  のときを調べればよい。補助の関数

$$I(\tau) = \int_D |\nabla \{e^{-\tau h_D(\omega)} v_\tau(x; \omega)\}|^2 dx$$

を導入する。(2.1) は

$$I(\tau) \leq I_\omega(\tau; h_D(\omega)) \leq C_D I(\tau) \quad (2.4)$$

と書かれる。 $\tau \rightarrow \infty$  のとき評価

$$I(\tau) = O(\tau^2) \quad (2.5)$$

をみるのは容易である。 $\partial D$  上の点  $x_0$  で直線  $x \cdot \omega = h_D(\omega)$  上にあるものを一つとる。 $\partial D$  は Lipschitz であると仮定しているから、 $x_0$  を頂点に持つ有限開錐  $C$  で  $D$  に含まれるものが存在する。このとき適当な正の定数  $a$  を取ると、 $s \rightarrow 0$  のとき  $C$  と直線  $x \cdot \omega = h_D(\omega) - s$  との共通部分の長さ  $|C \cap \{x \cdot \omega = h_D(\omega) - s\}|$  は  $as$  で与えられる。このとき十分大きい  $\tau_0(> 0)$ 、十分小さい  $s_0(> 0)$  を取ると、 $\tau_0$  以上のすべての  $\tau$  に対して  $I(\tau)$  の下からの評価

$$\begin{aligned} I(\tau) &\geq 2\tau^2 \int_C e^{-2\tau(h_D(\omega) - x \cdot \omega)} dx \\ &\geq 2\tau^2 \int_0^{s_0} e^{-2\tau s} |C \cap \{x \cdot \omega = h_D(\omega) - s\}| ds \\ &\geq a\tau^2 \int_0^{s_0} e^{-\tau s} s ds = a \int_0^{\tau s_0} e^{-s} s ds \\ &\geq a \int_0^{\tau_0 s_0} e^{-s} s ds \end{aligned} \quad (2.6)$$

を得る。(2.3) および他の主張はすべて (2.4), (2.5) および (2.6) からの帰結である。

□

この方法は、層状に重なった導体内に埋め込まれた介在物に対する類似の問題へ応用されている ([36, 38])。

## 2.2 単純化された囲い込み法

前部分節では無限に多くの  $f$  に対する  $\Lambda_D f$  を使って  $D$  の凸包を計算する公式を与えた。ここでは定数関数でない  $f$  を任意に与え固定し、このときの  $\Lambda_D f$  から、 $D$  が多角形状であるという仮定のもとで、その凸包を計算する公式を与えよう。その公式の確立に、Laplace 方程式の解の角における解析が重要な役割を果たしていることをみるであろう。

**定義 2.2.** 定数関数でない  $f \in H^{1/2}(\partial\Omega)$  を任意に与え固定する。与えられた  $\omega = (\omega_1, \omega_2) \in S^1$  および  $\tau(> 0)$  に対しもう一つの指示関数を公式

$$J_\omega(\tau; t) = \langle (\Lambda_0 - \Lambda_D)f, e^{-\tau t} v_\tau(\cdot; \omega) |_{\partial\Omega} \rangle \quad (2.7)$$

により定義する。

前部分節と異なりこの指示関数は複素数値である。

**定義 2.3.** 方向  $\omega \in S^1$  が  $D$  に関して正則であるとは、 $\partial D$  と直線  $x \cdot \omega = h_D(\omega)$  が一点のみを共有するときを言う。言い換えれば、方向  $\omega$  に垂直な直線  $x \cdot \omega = t$  が  $\omega$  方向の無限の彼方から下降してくるとき ( $t = \infty$  から  $t = -\infty$ )、 $t = h_D(\omega)$  のとき最初に  $\partial D$  にあたる。このとき、あたった点が一点のみであるというのが正則性の運動学的表現である。

**定義 2.4.**  $D$  が多角形状であるとは、 $D$  の各連結成分  $D_j$ ,  $j = 1, \dots, N$  がある多角形の内部で与えられるときを言う。

この部分節では次の定理を証明する。

**定理 2.3.**  $D$  は多角形状でかつ条件

$$\text{diam } D < \text{dist}(D, \partial\Omega) \quad (2.8)$$

を満たすと仮定する。方向  $\omega$  は  $D$  に関して正則であるとせよ。このとき公式

$$\lim_{\tau \rightarrow \infty} \frac{\log |J_\omega(\tau; 0)|}{\tau} = h_D(\omega) \quad (2.9)$$

が成り立つ。さらに以下の主張は正しい：

- もし  $t \geq h_D(\omega)$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} |J_\omega(\tau; t)| = 0;$$

- もし  $t < h_D(\omega)$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} |J_\omega(\tau; t)| = \infty.$$

与えられた多角形状の  $D$  に対して、それに正則でない方向は有限個である。したがって、任意に選んだ方向が  $D$  に関して正則でないことはめったにないということになる。さらに、(2.7) の  $f$  は定数関数でないという仮定以外、何の制限も課していないということとを特に強調しておく。

この定理の証明の出発点は、指示関数  $J_\omega(\tau; t)$  の、以下に述べる  $\partial D$  上における積分表現である。

命題 2.4.  $\lambda$  は任意の定数とせよ。 公式

$$J_\omega(\tau; t) = e^{-\tau t} \int_{\partial D} (u - \lambda) \frac{\partial}{\partial \nu} v_\tau(x; \omega) dS(x) \quad (2.10)$$

が成り立つ。

証明 本質は部分積分であり演習問題とする。

□

(2.10) の右辺を解析するためにある極座標を導入する。 $x_0$  で  $\partial D$  と直線  $x \cdot \omega = h_D(\omega)$  とのただ一つの共有点をあらわそう。 $x_0$  は  $D$  のある連結成分をその内部とする多角形の頂点でなければならず、このとき十分小さい正の数  $\eta$  をとると、 $x_0$  を中心とする半径  $2\eta$  の開円板と  $D$  の  $\Omega$  における外部  $\Omega \setminus \overline{D}$  との共通部分は扇形、 $\partial D$  と  $B(x_0, \eta)$  の共通部分は  $x_0$  で交わる二つの開線分と  $x_0$  からなる。 $\omega^\perp$  を  $x_0$  を中心として時計回りに回転すると、それら二つの線分に平行になるときがある。最初の線分に平行になったときの回転角を  $-p$  としその線分を  $\Gamma_p$  と書こう。つぎに二番目の線分に平行になったときの回転角を  $-q$  とし、その線分を  $\Gamma_q$  と書こう。すると  $0 < -p < -q < \pi$  すなわち  $-\pi < q < p < 0$  の関係があることがわかる。 $\Theta = 2\pi - \{(-q) - (-p)\}$  とおくと、 $\Theta + p = 2\pi + q$  および  $\pi < \Theta < 2\pi$  をみたすのは明らかである。この  $\Theta$  は上の扇形の中心角である。このとき

$$B(x_0, 2\eta) \cap (\Omega \setminus \overline{D}) = \{x_0 + r(\cos \theta \mathbf{a} + \sin \theta \mathbf{a}^\perp) \mid 0 < r < \eta, 0 < \theta < \Theta\}$$

および

$$\Gamma_p = \{x_0 + r(\cos p \omega^\perp + \sin p \omega) \mid 0 < r < \eta\}$$

$$\Gamma_q = \{x_0 + r(\cos q \omega^\perp + \sin q \omega) \mid 0 < r < \eta\}$$

と書けることがわかる。ここに

$$\mathbf{a} = \cos p \omega^\perp + \sin p \omega,$$

$$\mathbf{a}^\perp = -\sin p \omega^\perp + \cos p \omega.$$

$\mathbf{a}$  は  $\omega^\perp$  を時計回りに  $-p$  だけ回転して得られ、 $\mathbf{a}^\perp$  はそれをさらに反時計回りに  $\pi/2$  だけ回転して得られている。このとき

$$u(r, \theta) = u(x), \quad x = x_0 + r(\cos \theta \mathbf{a} + \sin \theta \mathbf{a}^\perp)$$

とおく。これが  $r = 0$  においてどんな特異性を持つか調べる。そのため定義域が

$$D(-\frac{d^2}{d\theta^2}) = \{\varphi \in H^2([0, \Theta]) \mid \varphi'(0) = \varphi'(\Theta) = 0\}$$

で与えられる  $L^2([0, \Theta])$  における自己共役作用素

$$-\frac{d^2}{d\theta^2}$$

を考える。数列  $\{\lambda_m\}$  および関数列  $\{\varphi_m\}$  を

$$\lambda_m = \frac{(m-1)\pi}{\Theta}, \quad m \geq 1;$$

$$\varphi_1(\theta) = \frac{1}{\sqrt{\Theta}};$$

$$\varphi_m(\theta) = \sqrt{\frac{2}{\Theta}} \cos \lambda_m \theta, \quad m \geq 2$$

により定義しよう。このとき  $\varphi_m, m = 1, \dots$  は  $-d^2/d\theta^2$  の正規化された固有関数系で、 $\lambda_m^2, m = 1, \dots$  は対応する固有値になっていることがわかる。次の命題は  $u(r, \theta)$  の  $r \rightarrow 0$  における挙動を明らかにする。

**命題 2.5.** 次の (1) から (3) が成り立つような実数からなる列  $\alpha_1, \alpha_2, \dots$  が存在する：

(1) 任意に固定した  $s \in ]0, 2[$  に対して

$$u(r, \theta) = \sum_{m=1}^{\infty} \alpha_m r^{\lambda_m} \varphi_m(\theta) \text{ in } H^1(B_{s\eta}(x_0) \cap (\Omega \setminus \overline{D})); \quad (2.11)$$

(2)  $m \rightarrow \infty$  のとき

$$|\alpha_m| = O\left(\frac{1}{\sqrt{\lambda_m}} \left(\frac{1}{2\eta}\right)^{\lambda_m}\right); \quad (2.12)$$

(3) 任意に固定した  $l = 1, 2, \dots$  に対して正の定数  $C_l$  が存在してすべての  $r \in ]0, \eta[$  に対して

$$\begin{aligned} |u(r, 0) - \sum_{m=1}^l \alpha_m r^{\lambda_m}| &\leq C_l r^{\lambda_{l+1}} \\ |u(r, \Theta) - \sum_{m=1}^l \alpha_m (-1)^{m-1} r^{\lambda_m}| &\leq C_l r^{\lambda_{l+1}}. \end{aligned} \quad (2.13)$$

証明は省略するが、[27] に Grisvard[21] に従った証明があることを注意しておく。この命題を使うと指示関数の  $\tau \rightarrow \infty$  のときの完全漸近展開を計算できる。まず  $\partial D \setminus B(x_0, \eta)$  が半平面  $x \cdot \omega \leq h_D(\omega) - \delta$  に含まれるように十分小さい  $\delta (> 0)$  をとり固定する。このとき (2.10) から評価

$$\begin{aligned} &J_\omega(\tau; h_D(\omega)) \\ &= e^{-\tau h_D(\omega)} \int_{\partial D \cap B(x_0, \eta)} \left\{ u - \frac{\alpha_1}{\sqrt{\Theta}} \right\} \frac{\partial}{\partial \nu} v_\tau(x; \omega) dS(x) + O(\tau e^{-\tau \delta}) \end{aligned} \quad (2.14)$$

を得る。

(2.13) と (2.14) を注意深く適用して次の定理を得る。

**定理 2.6.**  $\tau \rightarrow \infty$  のとき公式

$$J_\omega(\tau; h_D(\omega)) \sim -i \sqrt{\frac{2}{\Theta}} e^{i\tau x_0 \cdot \omega^\perp} \sum_{m=2}^{\infty} \frac{\Gamma(1 + \lambda_m) e^{i\frac{\pi}{2}\lambda_m} \alpha_m \{e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m}\}}{\tau^{\lambda_m}} \quad (2.15)$$

が成り立つ。

証明以下の等式は各自確認されたい：

$$\nu = (\sin p) \omega^\perp - (\cos p) \omega \text{ on } \Gamma_p$$

$$\nu = -(\sin q) \omega^\perp + (\cos q) \omega \text{ on } \Gamma_q$$

$$x \cdot \omega = h_D(\omega) + r \sin(\theta + p), \quad x \cdot \omega^\perp = x_0 \cdot \omega^\perp + r \cos(\theta + p)$$

$$\nabla v = \tau(\omega + i\omega^\perp) e^{\tau x \cdot (\omega + i\omega^\perp)}.$$

これらを使えば、等式

$$\begin{aligned} e^{-\tau h_D(\omega)} \frac{\partial v}{\partial \nu} &= -\tau e^{-ip} e^{i\tau x_0 \cdot \omega^\perp} e^{r\tau(\sin p + i \cos p)} \text{ on } \Gamma_p \\ e^{-\tau h_D(\omega)} \frac{\partial v}{\partial \nu} &= \tau e^{-iq} e^{i\tau x_0 \cdot \omega^\perp} e^{r\tau(\sin q + i \cos q)} \text{ on } \Gamma_q \end{aligned} \quad (2.16)$$

を得る。(2.13) と (2.16) から、任意に固定した  $l$  に対して  $\tau \rightarrow \infty$  のとき評価

$$|e^{-\tau h_D(\omega)} \int_{\Gamma_p} (u - \sum_{m=1}^l \alpha_m r^{\lambda_m}) \frac{\partial}{\partial \nu} v_\tau(x; \omega) dS| = O\left(\frac{1}{\tau^{\lambda_{l+1}}}\right) \quad (2.17)$$

$$|e^{-\tau h_D(\omega)} \int_{\Gamma_q} (u - \sum_{m=1}^l \alpha_m (-1)^{m-1} r^{\lambda_m}) \frac{\partial}{\partial \nu} v_\tau(x; \omega) dS| = O\left(\frac{1}{\tau^{\lambda_{l+1}}}\right)$$

を得る。(2.14)、(2.16) および (2.17) から漸近式

$$\begin{aligned} &J_\omega(t, h_D(\omega)) \\ &= -\sqrt{\frac{2}{\Theta}} e^{i\tau x_0 \cdot \omega^\perp} \sum_{m=2}^l \alpha_m \tau (e^{-ip} J_m(\tau, p) + (-1)^m e^{-iq} J_m(\tau, q)) + O\left(\frac{1}{\tau^{\lambda_{l+1}}}\right) \end{aligned} \quad (2.18)$$

を得る。ここで

$$J_m(\tau, \theta) = \int_0^\eta r^{\lambda_m} e^{r\tau(\sin \theta + i \cos \theta)} dr, \quad \theta = p, q$$

とおいた。これは

$$\begin{aligned} J_m(\tau, \theta) &= \tau^{-(1+\lambda_m)} \int_0^{\eta\tau} w^{\lambda_m} e^{w(\sin \theta + i \cos \theta)} dw \\ &= \tau^{-(1+\lambda_m)} \int_0^\infty w^{\lambda_m} e^{w(\sin \theta + i \cos \theta)} dw - \tau^{-(1+\lambda_m)} \int_{\eta\tau}^\infty w^{\lambda_m} e^{w(\sin \theta + i \cos \theta)} dw \end{aligned}$$

と書いて、公式

$$\int_0^\infty w^{\lambda_m} e^{w(\sin \theta + i \cos \theta)} dw = i e^{i\frac{\pi}{2}\lambda_m} e^{i\theta} e^{i\theta\lambda_m} \Gamma(1 + \lambda_m)$$

および評価

$$\int_{\eta\tau}^\infty w^{\lambda_m} e^{w(\sin \theta + i \cos \theta)} dw = O(\tau^{-\infty})$$

を使えば、 $\tau \rightarrow \infty$  のとき  $\tau$  に関して急減少の項を法として

$$\tau e^{-ip} J_m(\tau, \theta) = \frac{ie^{i\frac{\pi}{2}\lambda_m} e^{i\theta\lambda_m} \Gamma(1 + \lambda_m)}{\tau^{\lambda_m}} + O(\tau^{-\infty})$$

を得る。これと (2.18) から (2.15) が成り立つことが結論される。

□

この定理と自明な等式

$$J_\omega(\tau; t) = e^{-\tau(t-h_D(\omega))} J_\omega(\tau; h_D(\omega)) \quad \forall t \in \mathbf{R} \quad (2.19)$$

からただちに (2.9) が導かれるわけではない。(2.15) が保証するのは、 $J_\omega(\tau; h_D(\omega))$  が  $\tau \rightarrow \infty$  のとき高々代数的に減衰するというにすぎない。指数的減衰も高々代数的減衰であるから、このままでは  $t = 0$  または  $t < h_D(\omega)$  のときに何が起るのかまったくわからない。囲い込み法ではこの先を問題とする。 $J_\omega(\tau; h_D(\omega))$  は真に代数的減衰をする。それが次の補題の意味するところである。主張はなんの変哲もないようであるが証明は繊細である。

**補題 2.7.** すべての  $m \geq 2$  に対して

$$\alpha_m \{e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m}\} = 0$$

となることはない。

**証明**  $m$  を固定する。このとき

$$e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m} = 0$$

である必要十分条件は適当な整数  $k$  に対して

$$(2\pi - \Theta)\lambda_m = (2k + 1 + m)\pi \quad (2.20)$$

が成り立つことであるのは容易にみてとれる。

$$t = \frac{\Theta}{\pi}$$

とおこう。この  $t$  は  $1 < t < 2$  を満たす。このとき (2.20) は

$$m - 1 = t(k + m) \quad (2.21)$$

と書かれる。

まず  $e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m} = 0$  を満たすすべての  $m (\geq 2)$  からなる集合の構造を調べよう。最初に  $t$  が無理数である場合について考える。このときどんな  $m (\geq 2)$  に対しても (2.18) が成り立つ整数  $k$  は存在しない。したがってすべての  $m (\geq 2)$  に対して  $e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m} \neq 0$ 。

次に  $t$  が有理数である場合について考える。このとき

$$t = 1 + \frac{b}{a}$$

と書ける。ここに  $a$  は 2 以上の整数、 $b$  は 1 以上の整数でかつ  $a, b$  は互いに素である。さらに (2.21) は

$$(m + k)b + (k + 1)a = 0 \quad (2.22)$$

と書かれる。 $a, b$ は互いに素であるから、(2.22)は

$$m + k = la$$

がなりたつ整数  $l$  の存在を導く。これを (2.22) へ代入して

$$k = -(1 + lb)$$

と書かれることがわかる。このとき  $m = 1 + l(a + b)$  を得る。この逆も成り立つことは容易にわかる。

以上をまとめると次がわかったことになる。

$$\{m \geq 2 \mid e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m} = 0\} = \begin{cases} \emptyset, & t \text{ が無理数、} \\ \{1 + l(a + b) \mid l = 1, 2, \dots\}, & t \text{ が有理数。} \end{cases} \quad (2.23)$$

さてすべての  $m \geq 2$  に対して

$$\alpha_m \{e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m}\} = 0$$

が成り立つと仮定しよう。

$t$  が無理数であるとき。(2.23) からすべての  $m (\geq 2)$  に対して、 $\alpha_m = 0$  を得る。このとき命題 2.4 の (2.11) から  $u$  は  $B_{2\eta}(x_0) \cap (\Omega \setminus \overline{D})$  で定数関数であることがわかる。Laplace 方程式に対する一意接続定理によれば  $u$  は  $\Omega \setminus \overline{D}$  全体で定数関数でなければならない。

$t$  が有理数であるとき。(2.20) によれば、 $m \neq 1 + l(a + b)$ ,  $l = 1, 2, \dots$  ならば  $\alpha_m = 0$  を得る。再び (2.11) によればこのとき  $u$  は

$$u(r, \theta) = \sqrt{\frac{2}{\Theta}} \sum_{l=1}^{\infty} \alpha_{1+l(a+b)} r^{al} \cos(al\theta), \quad 0 < r < 2\eta, 0 < \theta < \Theta$$

と展開される。(2.12) を用いると、この右辺は  $B_{2\eta}(x_0)$  に含まれる  $\mathbf{R}^2$  の任意のコンパクト集合上その項別微分までこめて一様収束する。各項は  $\mathbf{R}^2$  全体で調和であるから  $B_{2\eta}(x_0)$  において調和である関数  $\tilde{u}(r, \theta)$  で

$$\begin{aligned} \tilde{u} &= u \text{ in } B_{2\eta}(x_0) \cap (\Omega \setminus \overline{D}) \\ \tilde{u}(r, \theta + \frac{2\pi}{a}) &= \tilde{u}(r, \theta) \end{aligned} \quad (2.24)$$

を満たすものが存在することを結論する。

$D$  は (2.8) を満たすと仮定しているから  $\text{diam } D < \delta < \text{dis}(D, \partial\Omega)$  を満たす正数  $\delta$  がとれる。このとき  $\overline{D} \subset B_\delta(x_0)$  および  $\overline{B}_\delta(x_0) \subset \Omega$  がなりたつことに注意する。あらかじめ  $\eta$  を十分小さくしておけば  $2\eta < \delta$  としてよい。 $x_0$  は  $x_0 \cdot \omega = h_D(\omega)$  を満足し  $\omega$  は  $D$  に関して正則であるから、 $\Omega \setminus \overline{D}$  に含まれる、 $x_0$  を中心とする半径  $d$  の二つの異なった半開円板が存在する。それらを取って固定する。 $S$  でそれら二つの半開円板と  $B_{2\eta}(x_0)$  との和集合をあらわし、 $S_j$  で  $S$  の  $x_0$  のまわりでの回転

$$\theta \mapsto \theta + \frac{2\pi(j-1)}{a}, \quad j = 1, \dots$$



による像をあらわそう。

ここで次の二つを注意する:

(1) 任意の  $j, k$  に対して  $S_j \cap S_k$  は連結である;

(2)  $B_d(x_0) = S_1 \cup S_2 \cup \cdots \cup S_a$ .

(1) は明らかであるが、しかし、(2) はそんなに自明であるとは思えない。ここでは念のためその証明の詳細を述べよう。

最初に  $x_0$  を中心とする別の極座標  $(r, \vartheta)$  で

$$S = B_{2\eta}(x_0) \cup \{(r, \vartheta) \mid -\pi\xi < \vartheta < 0, 2\eta \leq r < d\}$$

と  $S$  を表示するものをとる。ここに  $1 < \xi < 2$ 。このとき  $S_j$  は

$$S_j = B_{2\eta}(x_0) \cup \{(r, \vartheta) \mid -\pi\xi + \frac{2\pi}{a}(j-1) < \vartheta < \frac{2\pi}{a}(j-1), 2\eta \leq r < d\}$$

と書かれる。

$$S' = \{(r, \vartheta) \mid 0 \leq \vartheta \leq 2\pi - \pi\xi, 2\eta \leq r < d\}$$

とおこう。

$S_j, j = 1, \dots, a$  はすべて  $B_d(x_0)$  に含まれかつ  $B_d(x_0) = S_1 \cup S'$  が成り立つから上述の (2) の証明のためには

$$S' \subset S_1 \cup \cdots \cup S_a \quad (2.25)$$

を証明すれば十分である。

そこで  $(r, \vartheta) \in S'$  とせよ。 $\xi a > a \geq 2$  および  $\vartheta \leq 2\pi - \pi\xi$  が成り立つから

$$\frac{2\pi}{a}(a-1) - \vartheta \geq \frac{2\pi}{a}(a-1) - (2\pi - \pi\xi) = \frac{\pi}{a}(\xi a - 2) > 0$$

を得る。これは  $a$  が集合  $\{j = 1, \dots, a \mid \vartheta < \frac{2\pi}{a}(j-1)\}$  の要素であることを導き、したがってこの集合は空でないことがわかる。したがってその最小元

$$j_0 = \min\{j = 1, \dots, a \mid \vartheta < \frac{2\pi}{a}(j-1)\}$$

が存在する。 $j_0$  は  $1 \leq j_0 \leq a$  を満たす。 $j_0$  の最小性に注意すると、

$$\begin{aligned} \vartheta &\geq \frac{2\pi}{a}(j_0 - 2) \\ &= \frac{2\pi}{a}(j_0 - 1) - \frac{2\pi}{a} \\ &= -\pi\xi + \frac{2\pi}{a}(j_0 - 1) + \frac{\pi(a\xi - 2)}{a} \\ &> -\pi\xi + \frac{2\pi}{a}(j_0 - 1) \end{aligned}$$

を得る。これは  $(r, \vartheta) \in S_{j_0}$  であることを与えしたがって (2.25) は証明された。

さてこれ以降は、[19] で逆問題の解の一意性の証明で用いられた、(2.24) を満たす調和関数から出発する調和拡張のアイデアを適用する。まず  $S$  上で関数  $u'$  を以下のように定義する：

$$u'(x) = \begin{cases} \tilde{u}(x), & x \in B_{2\eta}(x_0), \\ u(x), & x \in S \setminus B_{2\eta}(x_0). \end{cases}$$

$u$  は  $\tilde{u}$  の調和拡張であるから  $u'$  が  $S$  で調和であることは明らかである。 $S_j$  上で関数  $u'_j$  を以下のように定義する：

$$u'_j(r, \theta) = u'(r, \theta - \frac{2\pi(j-1)}{a}) \text{ on } S_j.$$

$u'_j$  が  $S_j$  で調和であることは言うまでもない。

$B_{2\eta}(x_0) \subset S_j \cap S_k$  および (2.24) から任意の  $j, k$  に対して  $u'_j$  と  $u'_k$  は  $B_{2\eta}(x_0)$  上一致することがわかる。 $S_j \cap S_k$  は連結であるから、Laplace 方程式に対する一意接続定理により  $u'_j$  と  $u'_k$  は  $S_j \cap S_k$  上一致しなければならない。この性質は  $\cup_{j=1}^a S_j$  上で関数  $u''$  が

$$u'' = u'_j \text{ on } S_j, j = 1, \dots$$

によりうまく定義されることを導く。 $u''$  は調和でありかつ  $u$  と  $B_d(x_0) \setminus \overline{D}$  上一致する。 $\overline{D} \subset B_d(x_0) = S_1 \cup S_2 \cup \dots \cup S_a$  かつ  $B_d(x_0) \setminus \overline{D}$  は連結であるから、 $u''$  は  $u$  の  $B_d(x_0) \setminus \overline{D}$  から  $B_d(x_0)$  上への調和拡張になっている。

$\Omega$  上で関数  $u'''$  を

$$u'''(x) = \begin{cases} u''(x), & x \in B_d(x_0), \\ u(x), & x \in \Omega \setminus B_d(x_0) \end{cases}$$

により定義する。 $u'''$  は  $\Omega$  で調和かつ  $u$  と  $D$  の外部で一致する。いま、 $\overline{D}$  のある近傍で  $\Psi = 1$  を満たしかつ  $\text{supp } \Psi \subset B_d(x_0)$  である  $\Psi \in C_0^\infty(\Omega)$  をとる。(1.11) の弱解の定義から、

$$\begin{aligned} 0 &= \int_{\Omega \setminus \overline{D}} \nabla u \cdot \nabla(\Psi u''') dx \\ &= \int_{\Omega \setminus \overline{D}} \nabla u''' \cdot \nabla(\Psi u''') dx \\ &= \int_{\Omega} \nabla u''' \cdot \nabla(\Psi u''') dx - \int_D \nabla u''' \cdot \nabla(\Psi u''') dx \\ &= - \int_D \nabla u''' \cdot \nabla u''' dx \end{aligned}$$

を得る。したがって  $u'''$  は  $D$  の各連結成分上定数であることがわかる。再び一意接続定理を使えば  $f$  は定数でなければならないことが結論される。

□

補題 2.7 は最小元

$$m^* = \min\{m \geq 2 \mid \alpha_m \{e^{ip\lambda_m} + (-1)^m e^{iq\lambda_m}\} \neq 0\}$$

の存在を保証する。(2.15) の右辺は  $m = m^*$  から始まるから  $\tau \rightarrow \infty$  のとき

$$|J_\omega(\tau; h_D(\omega))| \tau^{\lambda_{m^*}} \rightarrow \sqrt{\frac{2}{\Theta}} \Gamma(1 + \lambda_{m^*}) |\alpha_{m^*} \{e^{ip\lambda_{m^*}} + (-1)^{m^*} e^{iq\lambda_{m^*}}\}| (> 0)$$

を得る。これと (2.19) から定理 2.3 のすべての主張を得る。

公式 (2.9) を基礎にした不連続性についての情報を抽出するためのアルゴリズムは [47] で提唱し、その数値的検証も行っている。後で分かったことであるが、**計算幾何学**という分野においては古くから有限個の方向に対する指示関数の誤差が混入した値から、物体の凸包の良い評価をいかに得るかという問題が研究されている ([60, 68])。 (2.9) はこの問題の一つの出所を提供したことになる。

最後に今後の問題について述べる。定理 2.2 は三次元においてもしかるべく形で成り立つことは容易にわかる (これは演習問題とする)。定理 2.3 は三次元においても  $D$  の各連結成分が、ある凸多面体の内部で与えられるときには成り立つと確信しているが、現在未解決である。静弾性論でも類似の問題が考えられるが特に興味があるのは定理 2.3 に相当する主張がこの場合になりつつかである。これも未解決である。このように、特に、単純化された囲い込み法の考え方は、**主張は単純だが証明となると骨のある興味ある問題**を産み出す。なお講演者は導電場の方程式 (1.9) の二次元版においても定理 2.3 に類似な定理を得ている ([31])。詳しくは、導電率  $\sigma$  が

$$\sigma(x) = \begin{cases} 1, & x \in \Omega \setminus D, \\ \sigma_1, & x \in D_1, \\ \vdots \\ \sigma_N, & x \in D_N \end{cases}$$

という形で与えられていて、 $\sigma_j$  がすべて 1 と異なる未知の正の定数であって、 $D$  が多角形状かつ (2.8) を満たすという仮定の下で、(2.9) に相当する公式が成り立つ。 $D$  は背景となる導体の導電率とは異なった導電率を持った導体であり、**介在物**に相当する。その公式を基礎にしたアルゴリズムの数値的検証は [48] ですでに報告済みである。

静弾性論でも類似の問題が考えられるが、これは未解決である。

無限に多くの電流密度分布に対応する、任意に固定した二点間の電位差を観測データとしてとった問題は [32] で考察されている。

層状に重なった導体内の層と層の間に発生した不完全接合部に対する類似の問題は層間剥離と関係していて興味ある問題であるが (その実験的研究は [62, 72] が興味深い)、[37] で二次元版を考察しその凸包を抽出する公式を与えた。3次元においては未解決である。

[39] では囲い込み法を音波の逆散乱問題へ応用した。補題 2.7 の議論とそこで得られた計算結果 ([39] の定理 4.1) を組み合わせると、**スペクトルの逆問題**について次の定理をただちに得ることができる。以下  $\Omega$  および  $D$  についての設定はこの節と同じである。 $k$  は正数で、 $u \in H^1(\Omega \setminus \overline{D})$  は Helmholtz 方程式

$$\Delta u + k^2 u = 0 \text{ in } \Omega \setminus \overline{D},$$

$$\frac{\partial u}{\partial \nu} = 0 \text{ on } \partial D$$

の**非自明解**であるとする。 $v_\tau$  に対応する Helmholtz 方程式の解として、関数  $\tilde{v}_\tau(x) = \tilde{v}_\tau(x; \omega) = e^{x \cdot (\tau \omega + i \sqrt{\tau^2 + k^2} \omega^\perp)}$  をとる。これと  $u$  の  $\partial\Omega$  上の Cauchy データを使って

$$\tilde{J}_\omega(\tau; t) = e^{-\tau t} \int_{\partial\Omega} \left( \frac{\partial u}{\partial \nu} \tilde{v}_\tau - \frac{\partial \tilde{v}_\tau}{\partial \nu} u \right) dS$$

とおこう。

**定理 2.8.**  $D$  は多角形状でかつ条件 (2.8) を満たすとする。 $k^2$  は、 $D$  のすくなくとも一つの連結成分における *Neumann* 境界条件の下での  $-\Delta$  の固有値ではないとする。方向  $\omega$  は  $D$  に関して正則であるとする。このとき定理 2.3 において  $J_\omega(\tau; t)$  を  $\tilde{J}_\omega(\tau; t)$  で置き換えた主張がなりたつ。

この定理の強力なところは、 $u$  の  $\partial\Omega$  における境界条件は何も課していないということと、ひとつの  $k$  だけしか使わないということである。著者はこの定理に相当する事実は弾性波に対しても成り立つと確信しているが、依然未解決である。

一般の楕円型方程式 (系) の角における特異性の解析自体は豊富な研究の蓄積 (例えば Grisvard [21]、Kozlov-Maz'ya-Rossmann[58]) がある。しかし、一般の楕円型方程式 (系) に対しても、囲い込み法の適用に必要なところまで知見が得られているのかどうか著者は知らない。専門家のご教示と協力を期待するところである。とはいえ、単純化された囲い込み法は、この種の研究に、逆問題の解の直接再構成への応用という新たな道を提供したことは確かであろう。

### 3 探針法と不連続性そのものの抽出

この節では空洞そのものをいかに観測データから抽出するかを問題とする。ここで紹介する探針法は、物体内に**仮想的な針**を侵入させてそれが未知の空洞に当たるか否かを、囲い込み法における指示関数に対応する**指示列**の挙動の違いによって捉えるものである。

#### 3.1. 針と針列

冒頭で述べた仮想的な針を定義しよう。

**定義 3.1.** 任意に与えられた  $\Omega$  内の点  $x \in \Omega$  に対して、以下の条件を満たす区分的に直線であるような曲線全体  $\sigma: [0, 1] \rightarrow \bar{\Omega}$  からなる集合を  $N_x$  であらわす：

- $\sigma(0) \in \partial\Omega$ ,  $\sigma(1) = x$  およびすべての  $t \in ]0, 1[$  に対して  $\sigma(t) \in \Omega$ ;
- $\sigma$  は写像として一対一。

$\sigma \in N_x$  を  $x$  を先端にもつ**針**と呼ぶ。

この針と空洞との位置関係をいかに Dirichlet-to-Neumann 写像を通して知るかが問題である。最初に  $\mathbf{R}^2$  における Laplace 方程式の基本解をとるのであるが、無限に多くある。探針法では基本解ならなんでも良いので標準的な基本解

$$G(y) = \frac{1}{2\pi} \log \frac{1}{|y|}$$

をとる。これを使って特別な調和関数列を定義する。

**定義 3.2.** 与えられた  $x \in \Omega$  および  $\sigma \in N_x$  に対して  $H^1(\Omega)$  の要素でかつ調和な関数からなる列  $\xi = \{v_n\}$  が  $(x, \sigma)$  に対する**針列**であるとは、 $\Omega \setminus \sigma([0, 1])$  に含まれる任意に固定した  $\mathbf{R}^2$  のコンパクト集合  $K$  に対して

$$\lim_{n \rightarrow \infty} (\|v_n(\cdot) - G(\cdot - x)\|_{L^2(K)} + \|\nabla\{v_n(\cdot) - G(\cdot - x)\}\|_{L^2(K)}) = 0$$

が成り立つときを言う。

そんな都合の良い調和関数列が存在するのかという疑問は当然であるが、それは Laplace 方程式に対する **Runge の近似定理**の帰結である。ここでは、ある作用素の像が稠密であることを示す論法でそれを構成しよう。証明で Hahn-Banach の拡張定理 (選択公理と同値、[70] の第三章) を使うのではとのご指摘を受けそうであるが、Hilbert 空間の枠内であ

れば、直交分解定理の使用におきかえられ、この点は避けられることを注意しておく ([70] の 65 ページ注意 3-3)。

**命題 3.1.** 任意に与えられた  $\Omega$  内の点  $x$  および  $x$  を先端にもつ針  $\sigma$  に対して、針列は存在する。

**証明論文** [43] の初期版で与えた証明を述べよう。 $\Omega \setminus \sigma([0, 1])$  に対して、その取り尽くし列  $\{O_n\}$  を以下のようにとれる：

- $\Omega \setminus \sigma([0, 1]) = \cup_n O_n$ ;
- 各  $O_n$  は滑らかな境界を持つ領域で、 $\Omega \setminus \overline{O_n}$  は連結；
- $\overline{O_n} \subset O_{n+1}$ .

$\overline{\Omega}$  を内部に含む開円板  $B$  をとり固定する。任意に与えられた  $g \in H^{-1/2}(\partial\Omega)$  に対して次の境界値問題の弱解  $u = u_g \in H_0^1(B)$  を考える：

$$\begin{aligned}\Delta u &= -T_g \text{ in } B, \\ u &= 0 \text{ on } \partial B.\end{aligned}\tag{3.1}$$

ここで  $T_g$  は

$$T_g(\varphi) = \langle g, \varphi|_{\partial\Omega} \rangle, \quad \varphi \in H_0^1(B)$$

で定義される  $H_0^1(B)$  上の有界線形汎関数である； $H_0^1(B)$  は  $H_1(B)$  の要素でその  $\partial B$  上へのトレースが消えるようなもの全体からなる、 $H_1(B)$  の閉部分空間をあらわす。各  $n$  に対して

$$K_n g = u_g|_{\partial O_n} \in H^{1/2}(\partial O_n)$$

により  $H^{-1/2}(\partial\Omega)$  から  $H^{1/2}(\partial O_n)$  の中への有界線形作用素  $K_n$  を定義する。

このとき集合

$$\{K_n g \mid g \in L^2(\partial\Omega)\}$$

は  $H^{1/2}(\partial O_n)$  で稠密であることを証明しよう。

Hilbert 空間  $H^{1/2}(\partial O_n)$  における **Riesz の表現定理** と **直交分解定理** によれば  $T \in H^{1/2}(\partial O_n)$  が、すべての  $g \in L^2(\partial\Omega)$  に対して

$$(K_n g, T) = 0$$

を満たすならば、そのときすべての  $h \in H^{1/2}(\partial O_n)$  に対して

$$(h, T) = 0$$

が成り立つことを証明することが必要十分である。ここで  $(, )$  は  $H^{1/2}(\partial O_n)$  における内積である。

$H_0^1(B)$  上の有界線形汎関数  $\tilde{T}$  を

$$\tilde{T}(\varphi) = (\varphi|_{\partial O_n}, T), \quad \varphi \in H_0^1(B)$$

により定義する。さらに  $\tilde{u} \in H_0^1(B)$  で次の境界値問題の弱解をあらわそう：

$$\Delta \tilde{u} = -\tilde{T} \text{ in } B,$$

$$\tilde{u} = 0 \text{ on } \partial B.$$

そのとき仮定より

$$\begin{aligned}
0 &= (K_n g, T) \\
&= (u_g|_{\partial O_n}, T) = \tilde{T}(u_g) \\
&= \int_B \nabla \tilde{u} \cdot \nabla u_g dx \\
&= \int_B \nabla u_g \cdot \nabla \tilde{u} dx \\
&= T_g(\tilde{u}) \\
&= \langle g, \tilde{u}|_{\partial \Omega} \rangle \\
&= \int_{\partial \Omega} g \tilde{u}|_{\partial \Omega} dS.
\end{aligned}$$

$g$  は  $L^2(\partial \Omega)$  の任意の要素であるので、 $\tilde{u}$  は  $\partial \Omega$  上消えていなければならない。関数  $\tilde{u}$  は  $B \setminus \partial O_n$  で調和で  $\partial B$  上  $\tilde{u} = 0$  であるから、 $B \setminus \bar{\Omega}$  における Dirichlet 問題の解の一意性により  $\tilde{u}$  は  $B \setminus \bar{\Omega}$  全体で消えていることがわかる。そのとき、Laplace 方程式に対する一意接続定理および  $\Omega \setminus \bar{O}_n$  の連結性から  $\tilde{u}$  は  $\Omega \setminus \bar{O}_n$  で消えなければならない、したがって  $\partial O_n$  上  $\tilde{u} = 0$  を得る。

任意に与えられた  $h \in H^{1/2}(\partial O_n)$  に対して関数  $\tilde{v} \in H_0^1(B)$  を次のようにとる：

- $\partial O_n$  上  $\tilde{v} = h$ ；
- $\tilde{v}$  は  $O_n$  で調和。

そのとき目的の結論を以下のように得る：

$$\begin{aligned}
(h, T) &= (\tilde{v}|_{\partial O_n}, T) \\
&= \tilde{T}(\tilde{v}) \\
&= \int_B \nabla \tilde{u} \cdot \nabla \tilde{v} dx \\
&= \int_{O_n} \nabla \tilde{v} \cdot \nabla \tilde{u} dx = 0.
\end{aligned}$$

正数からなる列  $\epsilon_1, \epsilon_2, \dots$  を任意に与える。各  $n$  に対して  $G(\cdot - x)|_{O_n}$  は  $H^1(O_n)$  に属するから、上の事実により、 $L^2(\partial \Omega)$  の要素  $g_n = g_n(\cdot; x)$  で評価

$$\|K_n g(\cdot; x) - G(\cdot - x)|_{\partial O_n}\|_{H^{1/2}(\partial O_n)} < \epsilon_n \quad (3.2)$$

を満たすものが存在する。このとき  $g = g_n(\cdot; x)$  に対する (3.1) の弱解  $u = u_n(\cdot; x)$  を考える。 $u$  および  $G(\cdot; x)$  とともに  $O_n$  で調和で、 $O_n$  における Laplace 方程式に対する Dirichlet 境界値問題の適切性と (3.2) を組み合わせれば、評価

$$\|u - G(\cdot - x)\|_{H^1(O_n)} \leq C_n \epsilon_n$$

を得る。ここで  $C_n$  は  $O_n$  のみに依存する正定数である。そこで  $\epsilon_n$  を  $n \rightarrow \infty$  のとき

$$C_n \epsilon_n \rightarrow 0$$

を満たすように選んでやれば、 $v_n = u_n|_{\Omega}$  とおくと、 $\xi = \{v_n\}$  は  $(x, \sigma)$  に対する針列である。

□

$\epsilon_n$  を具体的に与えるためには、 $C_n$  の  $n \rightarrow \infty$  における増大度を知ることは大変重要である。どれくらいであろうか。

ここで定義した針列は後でみるように、まさに針のような役割を演じる。

**定義 3.3.** 与えられた  $\Omega$  内の点  $x$ 、 $x$  を先端に持つ針  $\sigma$  および  $(x, \sigma)$  に対する針列  $\xi = \{v_n\}$  に対して数列

$$I(x, \sigma, \xi)_n = \langle (\Lambda_0 - \Lambda_D) f_n, \bar{f}_n \rangle, \quad n = 1, 2, \dots$$

を指示列と言う。ここに  $f_n$  は  $v_n$  の  $\partial\Omega$  上へのトレースである。

ここで探針法とはなにかはっきり言うことができる。探針法とは多くの針に対する指示列の挙動を監視することによって  $\Omega$  の内部を探索する方法である。

### 3.2. 探針法の A 面

指示列の  $n \rightarrow \infty$  のときの挙動は二つの面をもつ。ここではその一つの面を説明する。 $\Omega$  内の与えられた点  $x$  および  $\sigma \in N_x$  に対して  $\xi = \{v_n\}$  を  $(x, \sigma)$  に対する任意の針列とする。(2.2) から指示列の積分表現

$$I(x, \sigma, \xi)_n = \int_{\Omega \setminus \bar{D}} |\nabla w_n|^2 dy + \int_D |\nabla v_n|^2 dy$$

を得る。ここに  $w_n = u_n - v_n$  であり、 $u_n \in H^1(\Omega \setminus \bar{D})$  は (1.11) の  $f = v_n|_{\partial\Omega}$  に対する弱解である。関数  $w_n$  は次の境界値問題

$$\Delta w = 0 \text{ in } \Omega \setminus \bar{D},$$

$$\frac{\partial w}{\partial \nu} = -\frac{\partial v_n}{\partial \nu} \text{ on } \partial D,$$

$$w = 0 \text{ on } \partial\Omega$$

のしかるべく定義された弱解になっている。この講義では  $w_n$  を  $v_n$  の空洞  $D$  による反射解と呼ぶ。

ここで、 $x$  は  $D$  の  $\Omega$  における外部  $\Omega \setminus \bar{D}$  にありかつ  $\sigma \in N_x$  は  $\bar{D}$  と共有点を持たないとする。これは導体の表面から針をさしてそれがまだ空洞の表面に到達していない状況に対応する。このとき針列はその定義から  $D$  上  $H^1(D)$  の位相で  $G(\cdot - x)$  に収束する。(1.14) を使うと反射解の列  $\{w_n\}$  はある関数  $w_x$  に  $n \rightarrow \infty$  のとき  $H^1(\Omega \setminus \bar{D})$  の位相で収束する。 $w_x$  は次の境界値問題のしかるべく定義された弱解である：

$$\Delta w = 0 \text{ in } \Omega \setminus \bar{D},$$

$$\frac{\partial w}{\partial \nu} = -\frac{\partial}{\partial \nu}(G(\cdot - x)) \text{ on } \partial D,$$

$$w = 0 \text{ on } \partial\Omega.$$

この  $x \in \Omega \setminus \bar{D}$  をパラメタに持つ  $w_x$  を空洞  $D$  による反射解と呼ぶ。これらの事実から指示列  $\{I(x, \sigma, \xi)_n\}$  は  $n \rightarrow \infty$  のとき次で定義される関数に収束することがわかる。

**定義 3.4.** 探針法における 指示関数  $I$  を

$$I(x) = \int_{\Omega \setminus \overline{D}} |\nabla w_x|^2 dy + \int_D |\nabla G(y-x)|^2 dy, \quad x \in \Omega \setminus \overline{D}$$

により定義する。ここで  $w_x$  は空洞  $D$  による反射解である。

次に述べる定理は探針法の一つの側面 (A 面と呼ぶ) を表現している。それは、指示関数の空洞の表面における爆発とその指示関数を導体の表面から適当な針に対する指示列を使って計算する方法である。

**定理 3.2.A.**

- (A.1) 任意に与えられた  $\Omega \setminus \overline{D}$  内の点  $x$  および  $x$  を先端に持つ針  $\sigma$  に対してもし  $\sigma([0, 1])$  が  $\overline{D}$  と共有点を持たないならば、そのとき  $(x, \sigma)$  に対する任意の針列  $\xi = \{v_n\}$  に対して指示列  $\{I(x, \sigma, \xi)_n\}$  は指示関数  $I(x)$  に収束する;
- (A.2) 任意に固定した正数  $\epsilon$  に対して

$$\sup_{\text{dist}(x, \overline{D}) > \epsilon} I(x) < \infty$$

が成り立つ;

- (A.3)  $\partial D$  上の任意に与えられた点  $a$  に対して

$$\lim_{x \rightarrow a} I(x) = \infty$$

が成り立つ。

**証明** (A.1) および (A.2) はこの部分説の冒頭で述べたことの帰結である。(A.3) は自明な不等式

$$I(x) \geq \int_D |\nabla G(y-x)|^2 dy, \quad x \in \Omega \setminus \overline{D}$$

および事実 (演習問題とする)

$$\lim_{x \rightarrow a} \int_D |\nabla G(y-x)|^2 dy = \infty$$

からの帰結である。

□

(A.2) および (A.3) によれば指示関数はその値を考える点が空洞に近づくときと大きくなる量である。そこで導体の表面から針を少し侵入させ適当な針列を使って指示関数のその先端における値を (A.1) を使って計算する。そしてさらに針を伸ばし同じ計算をしていく。これ続けると針の侵入経路上に空洞があれば、針の先端が空洞に近づくに従って計算された量はやがて大きくなるであろう。そうでなければ大きくはならないであろう。これが探針法の A 面から示唆される空洞の位置を探る方法ある。言うのは簡単であるが実際にやるにはいろいろ解決すべき点があるのはあきらかである。針列の構成はどうか。理論的には、針に依存した、くびれた領域における第一種の積分方程式の最小ノルム解 [59] の数値的構成に帰着できることは分かっていたが、最近 Cheng-Liu-Nakamura[10] はこの数値的構成をし、A 面から示唆される方法の数値実験を実行し、他の間接再構成方法のための初期の推測を与えるであろうという知見を得ている。Erhard-Potthast[18] は、針列のより粗い数値的構成を行い、同様の数値実験を行っているが、興味ある結果を得ている。

### 3.3. 探針法の B 面



定理 3.2.A の (A.1) を見れば誰でも次のような疑問を抱くだろう。もし針の先端がちょうど空洞の表面上に到達したときまたは空洞の内部に侵入したとき、さらには空洞を貫通してしまったとき、指示列は  $n \rightarrow \infty$  のときどんな挙動をするか。これは探針法の初期の段階では不問に付されていた問題で、最近解決された。それが次の定理である。

**定理 3.2.B.**  $x$  は  $\Omega$  内の任意の点、 $\sigma$  は  $x$  を先端に持つ針とする。もし  $x$  が  $\Omega \setminus \overline{D}$  内の点で  $\sigma([0, 1])$  と  $D$  が共有点をもつかまたは  $x$  が  $\overline{D}$  内の点であるならば、そのとき  $(x, \sigma)$  に対する任意の針列  $\xi = \{v_n\}$  に対して  $\lim_{n \rightarrow \infty} I(x, \sigma, \xi)_n = \infty$  が成り立つ。

この証明の核心は、針列の針上での爆発を主張する以下の二つの補題である。

最初の補題は、針列の、針の先端におけるエネルギーの爆発を主張している。

**補題 3.3.**  $x$  は  $\Omega$  内の任意の点および  $\sigma$  は  $x$  を先端に持つ針とする。 $\xi = \{v_n\}$  を  $(x, \sigma)$  に対する任意の針列とする。このとき  $x$  を頂点にもつ任意の有限開錐  $V$  に対して

$$\lim_{n \rightarrow \infty} \int_{V \cap \Omega} |\nabla v_n(y)|^2 dy = \infty$$

が成り立つ。

証明背理法による。結論が成り立たないと仮定する。このとき正数  $M$  および自然数列  $n_1 < n_2 < \dots \rightarrow \infty$  で

$$\int_{V \cap \Omega} |\nabla v_{n_j}(y)|^2 dy < M, j = 1, 2, \dots$$

を満たすものが存在する。十分小さい半径  $R$  を持つ  $x$  を中心とする開円板  $B$  をとると  $\overline{B}$  は  $\Omega$  に含まれかつ  $\sigma([0, 1])$  と  $B$  との共通部分は  $x$  をその一端とする線分になる。このとき、 $V$  に含まれる、 $x$  を頂点に持つ有限開錐  $V'$  で、 $0 < \epsilon < R$  を満たす任意の  $\epsilon$  に対して  $K_\epsilon \equiv \overline{V'} \cap (\overline{B} \setminus B_x(\epsilon))$  が  $V \cap (\Omega \setminus \sigma([0, 1]))$  に含まれるものが存在する。ここで  $B_x(\epsilon)$  は  $x$  を中心とする半径  $\epsilon$  の開円板をあらわす。したがって

$$\int_{K_\epsilon} |\nabla v_{n_j}(y)|^2 dy < M, j = 1, 2, \dots$$

を得る。 $\nabla v_{n_j}$  は  $\nabla G(\cdot - x)$  に  $L^2(K_\epsilon)$  の位相で収束するから上の不等式から

$$\int_{K_\epsilon} |\nabla G(y - x)|^2 dy \leq M$$

を得る。 $\epsilon$  はいくらでも小さくできるから  $\epsilon = 1/l$  に対して **Fatou の補題** を適用して

$$\int_{V' \cap B} |\nabla G(y - x)|^2 dy \leq M$$

を得る。しかし、 $x$  を中心とする極座標を用いればわかるように、この左辺は発散する。これは矛盾である。

□

もう一つの補題は、針列の、針の中間点でのエネルギーの爆発の主張であるが、それを述べる前に大変有用な評価を述べる。

**命題 3.4.**  $W$  は  $\mathbb{R}^2$  のその境界が  $C^2$  である任意の有界領域であるとする。 $v \in H^1(W)$  は  $W$  において Laplace 方程式  $\Delta v = 0$  を満たすとせよ。このとき  $W$  に含まれる任意の  $\mathbb{R}^2$  のコンパクト集合  $K$  に対して  $v$  に独立な正の定数  $C' = C'(K, W)$  が存在して評価

$$\int_K |\nabla v|^2 dy \leq C' \int_W |v|^2 dy \quad (3.3)$$

が成り立つ。さらに  $v$  に独立な正の定数  $C = C(W)$  が存在して評価

$$\int_W |v|^2 dy \leq C \int_{\partial W} |v|^2 dS \quad (3.4)$$

が成り立つ。

証明最初に (3.3) の証明を述べる。 $\varphi \in C_0^\infty(W, \mathbf{R})$  を任意に取る。 $W$  における方程式  $\Delta v = 0$  の両辺に  $\varphi^2 \bar{v}$  をかけてその結果を  $W$  上で積分する。部分積分を使って不等式

$$\int_W |\nabla v|^2 \varphi^2 dy \leq C \int_W |v|^2 |\nabla \varphi|^2 dy$$

を得る。ここで  $\varphi \in C_0^\infty(W)$  を、 $K$  の上で恒等的に 1 をとり  $0 \leq \varphi \leq 1$  を満たすように取る。このときただちに (3.3) を得る。

次に  $z \in H^2(W)$  は次の境界値問題

$$\Delta z = \bar{v} \text{ in } W,$$

$$z = 0 \text{ on } \partial W$$

を解いて得られる関数とせよ。ここは詳しくいうと、まずその  $\partial W$  上へのトレースが消えるような  $H^1(W)$  の中の要素としてこの問題の弱解を求め、つぎにその境界まで込めた正則性を示すことで  $z \in H^2(W)$  および評価

$$\|z\|_{H^2(W)} \leq C_1 \|v\|_{L^2(W)}$$

を得る。ここに  $C_1 = C_1(W) > 0$ 。このとき

$$\begin{aligned} \int_W |v|^2 dy &= \int_W (\Delta z) v dy \\ &= \int_{\partial W} \frac{\partial z}{\partial \nu} v dS - \int_W \nabla z \cdot \nabla v dy \\ &= \int_{\partial W} \frac{\partial z}{\partial \nu} v dS \end{aligned}$$

およびトレース定理から

$$\begin{aligned} \|v\|_{L^2(W)}^2 &\leq \int_{\partial W} \left| \frac{\partial z}{\partial \nu} \right| |v| dS \\ &\leq \|\nabla z\|_{L^2(\partial W)} \|v\|_{\partial W} \|v\|_{L^2(\partial W)} \\ &\leq C_2 \|z\|_{H^2(W)} \|v\|_{\partial W} \|v\|_{L^2(\partial W)} \\ &\leq C_1 C_2 \|v\|_{L^2(W)} \|v\|_{\partial W} \|v\|_{L^2(\partial W)} \end{aligned}$$

を得る。ここで  $C_2 = C_2(W) > 0$  はトレース定理からくる定数である。これは (3.4) を導く。

□

二つめの補題は次のようである。

補題 3.5.  $x$  は  $\Omega$  内の任意の点および  $\sigma$  は  $x$  を先端に持つ針とする。  $\xi = \{v_n\}$  を  $(x, \sigma)$  に対する任意の針列とする。このとき  $\sigma([0, 1])$  上の任意の点  $z$  および  $z$  を中心とする任意の開円板  $B$  に対して

$$\lim_{n \rightarrow \infty} \int_{B \cap \Omega} |\nabla v_n(y)|^2 dy = \infty$$

が成り立つ。

証明  $v$  は  $\Omega$  における任意の調和関数とする。まず粗く言って針の先端における  $\nabla v$  を使って同じ針の中間点における  $\nabla v$  を下から評価することを考える。 $v$  は  $\Omega$  における滑らかな関数と同一視でき、そのすべての偏導関数も調和である。 $z$  を中心とする開円板  $B'$  でその閉包が  $B \cap \Omega$  に含まれるものをとる。次に同じく  $z$  を中心とするより小さい半径を持つ開円板  $B''$  をとる。 $W = B'$  および  $K = \overline{B''}$  の場合における (3.3) を使うと評価

$$\int_{B''} |\nabla(\nabla v)|^2 dy \leq C \int_{B'} |\nabla v|^2 dy \quad (3.5)$$

を得る。 $B''$  におけるトレース定理によれば

$$\int_{\partial B''} |\nabla v|^2 dS \leq C' \left( \int_{B''} |\nabla v|^2 dy + \int_{B''} |\nabla(\nabla v)|^2 dy \right) \quad (3.6)$$

が成り立つ。 $\sigma$  は区分的に直線であるから、 $x$  を含む  $C^2$  領域  $U$  を、その境界と  $\partial B''$  との共通部分  $\Sigma \equiv \partial U \cap \partial B''$  が  $\partial B''$  上、正の長さを持ちかつ、 $\text{dist}(\partial U \setminus \Sigma, \sigma) > 0$  を満たすようにとれるのは明らかである。

さて  $x$  を中心とする開円板  $B'''$  でその閉包が  $U$  に含まれるものをとる。 $W = U$  の場合における (3.4) を使うと評価

$$\begin{aligned} & \int_{B'''} |\nabla v|^2 dy \\ & \leq \int_U |\nabla v|^2 dy \\ & \leq C'' \int_{\partial U} |\nabla v|^2 dS \\ & = C'' \left( \int_{\Sigma} |\nabla v|^2 dS + \int_{\partial U \setminus \Sigma} |\nabla v|^2 dS \right) \\ & \leq C'' \left( \int_{\partial B''} |\nabla v|^2 dS + \int_{\partial U \setminus \Sigma} |\nabla v|^2 dS \right) \end{aligned} \quad (3.7)$$

を得る。(3.5)、(3.6) および (3.7) を使うと、 $B'$  における  $\nabla v$  を  $B'''$  における  $\nabla v$  を使って次のように評価できる：

$$\int_{B'''} |\nabla v|^2 dy \leq C''' \left( \int_{\partial U \setminus \Sigma} |\nabla v|^2 dS + \int_{B'} |\nabla v|^2 dy \right). \quad (3.8)$$

さて  $v = v_n(\cdot)$  としよう。 $\nabla v_n(\cdot)$  が  $\nabla G(\cdot - x)$  に  $H_{\text{loc}}^1(\Omega \setminus \sigma([0, 1]))$  の位相で収束するのは (3.5) を適当な  $W$  に適用してみてもとれる。 $\partial U \setminus \Sigma$  は  $\sigma([0, 1])$  から正の距離をもって離れているからトレース定理により

$$\lim_{n \rightarrow \infty} \int_{\partial U \setminus \Sigma} |\nabla v_n(y)|^2 dy = \int_{\partial U \setminus \Sigma} |\nabla G(y - x)|^2 dy < \infty \quad (3.9)$$

を結論する。一方補題 3.3 によれば

$$\lim_{n \rightarrow \infty} \int_{B'''} |\nabla v_n(y)|^2 dy = \infty \quad (3.10)$$

がわかっている。したがって  $v = v_n(\cdot)$  に対する (3.8)、(3.9) および (3.10) より補題は証明された。

□

定理 3.2.B の証明を述べよう。  $I(x, \sigma, \xi)_n$  の発散は数列

$$\int_D |\nabla v_n|^2 dy \quad (3.11)$$

の発散からただちに従うことは言うまでもない。もし  $x$  が  $D$  内にあれば、補題 3.3 から (3.11) の発散を得る。もし  $x$  が  $\partial D$  上にあれば、 $x$  を頂点に持つ有限開錐  $V$  で  $D$  に含まれるものがとれる。これは  $\partial D$  の Lipschitz 正則性から結論される。再び補題 3.3 は (3.11) の発散を導く。 $x$  が  $\Omega \setminus \overline{D}$  の要素でかつ  $\sigma([0, 1]) \cap D \neq \emptyset$  のときは補題 3.5 を使えばよい。

探針法の二つの面をひとつにまとめると、未知の空洞を指示列の挙動を使って特徴づけることができる。それが次の系である。

**系 3.6.** 与えられた  $\Omega$  内の点  $x$  が  $\Omega \setminus \overline{D}$  の要素であるための必要十分条件は  $x$  を先端に持つ針  $\sigma$  および  $(x, \sigma)$  に対する針列  $\xi$  で指示列  $\{I(x, \sigma, \xi)_n\}$  が上に有界であるものが存在することである。

証明は演習問題とする。

探針法の A 面はより複雑な問題においても適用できている。[35] では弾性体中の介在物を抽出する問題、[16] では導体の境界条件が混合型である場合が考察されている。また [9] では支配方程式が Helmholtz で  $D$  の境界条件が Robin である場合が取り上げられた。著者はプレプリント [44] で同じ問題を取り上げ、別証明を与えると同時に、B 面について、ある条件のもとで、肯定的結果を与えた。

## 4 囲い込み法と Mittag-Leffler の関数

この節では二節で紹介した囲い込み法の、その精神を生かした一般化について述べる。ここであえて一般化という言葉を使ったが、これはその内容を薄めてより広い問題へ適用できるようにするという意味ではない。真の一般化とは、以前と同じ問題にその一般化の考え方を適用したときに、以前に得られた結果以上の知見が得られるとき言う。この意味で、ここで展開される考え方は囲い込み法の一般化である。

ではなぜ囲い込み法の一般化を考える必要があるのか。それは、もし空洞の連結成分が導体内に広く散らばっている場合、その凸包は導体の広い部分を占めることになりその詳細が隠れてしまうからである。もうひとつは数学的な興味である。囲い込み法は支配方程式の指数関数解の族の性質を基礎にしているが、凸包以上の情報をもたらす、探針法における針列とは違った、はっきり計算できる解がないであろうか。

この疑問に対する一つの答えを展開するのがこの節である。

### 4.1. 囲い込み法の一般化

複素変数  $z$  の整関数

$$E_\alpha(z) = \sum_{m=0}^{\infty} \frac{z^m}{\Gamma(\alpha m + 1)}$$

は Mittag-Leffler の関数と呼ばれる。 $\alpha$  は正のパラメタで、ここでは、その動く範囲を  $\alpha \leq 1$  に制限しておく。 $E_1(z) = e^z$  であり、したがって Mittag-Leffler の関数は指数関数をその特別な場合として含んでいる。この節ではこれから以降は  $\alpha < 1$  の場合のみ考える。このとき  $E_\alpha(z)$  は、 $|z| \rightarrow \infty$  のとき、次の注目すべき漸近挙動を示すことが知られている。すなわち、原点を頂点にもつある角領域にとどまりつつ無限遠に向かう場合は絶対値が (指数的に) 増大し、その外部にとどまりつつ無限遠に向かう場合は絶対値は (代数的に) 減衰する。しかも角領域の開き角度はパラメタ  $\alpha$  を小さくすればするほど鋭くできる。この詳細については部分節 4.2 の補題 4.3 をみられたい。

簡単に言うと、一般化された囲い込み法とは、Mittag-Leffler の関数のこの性質を利用して、未知の空洞の凸包以上の情報を抽出する方法である。では凸包以上とはどういう意味か。それをはっきり述べると次のようになる。

**定義 4.1.** 任意に与えられた  $\Omega$  内の点が **可視** であるとは、その点から無限遠点まで  $\bar{D}$  に触れることなくある直線に沿って移動できるときを言う。可視である点全体からなる集合を  $V(D)$  であらわす。

一般化された囲い込み法はこの  $V(D)$  を  $\Lambda_D$  から抽出する。

すぐわかることは、 $\bar{D}$  内にあるどの点も可視でない。したがって  $V(D)$  を知ることは次の意味での  $\bar{D}$  の上からの評価をもたらす：

$$\bar{D} \subset \Omega \setminus V(D).$$

一般には等式  $\bar{D} = \Omega \setminus V(D)$  は成り立たない。できるだけ複雑な成り立つ例またはできるだけ簡単な成り立たない例を与えることは演習問題とする。

さて囲い込み法における調和関数  $v_\tau(x; \omega) = e^{\tau x \cdot (\omega + i\omega^\perp)}$  の代わりを演じる調和関数は、与えられた点  $y \in \mathbf{R}^2$  および  $\tau > 0$  に対して

$$e_\tau^\alpha(x; y, \omega) = E_\alpha(\tau(x - y) \cdot (\omega + i\omega^\perp)), \quad x \in \mathbf{R}^2$$

と定義される関数である。この調和関数を用いて一般化された囲い込み法における指示関数を導入する。

**定義 4.2.** 一般化された囲い込み法における **指示関数** を

$$I_{(y, \omega)}^\alpha(\tau) = \langle (\Lambda_0 - \Lambda_D)(e_\tau^\alpha(\cdot; y, \omega)|_{\partial\Omega}), \overline{e_\tau^\alpha(\cdot; y, \omega)|_{\partial\Omega}} \rangle$$

により定義する。

この指示関数は  $\omega$  に直交する二つの単位ベクトルのどちらを  $\omega^\perp$  としても変わらない。そこでここでは、第二節とは向きを反対にして  $\omega^\perp = (-\omega_2, \omega_1)$  ととる。これは単に記述の便利のためである。したがって  $\omega, \omega^\perp$  の向きはこの順に  $\mathbf{R}^2$  の標準的な向きに一致する。与えられた点  $y$  および方向  $\omega$  に対して  $\omega$  の回りの開き角が  $\pi\alpha/2$  である  $y$  を頂点に持つ開錐を  $C_y(\omega, \pi\alpha/2)$  であらわす：

$$C_y(\omega, \pi\alpha/2) = \{x \in \mathbf{R}^2 \mid (x - y) \cdot \omega > |x - y| \cos(\pi\alpha/2)\}.$$

このとき、点  $y$  が可視であるのは、適当な  $\alpha \in ]0, 1[$  および  $\omega \in S^1$  に対して  $\{C_y(\omega, \pi\alpha/2)\} \cap \bar{D} = \emptyset$  がなりたつきおよびそのときに限るのは明らかである。

次の定理は、指示関数の  $\tau \rightarrow \infty$  における漸近挙動から、与えられた錐  $C_y(\omega, \pi\alpha/2)$  が  $\bar{D}$  に触れるかどうか分かるということを述べている。

**定理 4.1.** 任意に与えられた  $\Omega$  内の点  $y$  および方向  $\omega$  に対して、指示関数  $I_{(y, \omega)}^\alpha(\tau)$  の  $\tau \rightarrow \infty$  のときの漸近挙動について以下が成り立つ：

- もし  $\overline{\{C_y(\omega, \pi\alpha/2)\}} \cap \overline{D} = \emptyset$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} |I_{(y, \omega)}^\alpha(\tau)| = 0;$$

- もし  $\overline{\{C_y(\omega, \pi\alpha/2)\}} \cap \overline{D} \neq \emptyset$  および  $C_y(\omega, \pi\alpha/2) \cap D = \emptyset$  ならば、そのとき

$$\liminf_{\tau \rightarrow \infty} |I_{(y, \omega)}^\alpha(\tau)| > 0;$$

- もし  $C_y(\omega, \pi\alpha/2) \cap D \neq \emptyset$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} |I_{(y, \omega)}^\alpha(\tau)| = \infty.$$

この系として、 $V(D)$  内の点  $y$  の、たくさんの  $\omega$  および  $\alpha \in ]0, 1[$  に対する  $I_{(y, \omega)}^\alpha(\tau)$  の  $\tau \rightarrow \infty$  における漸近挙動による特徴づけを得る。

#### 系 4.2. 公式

$$V(D) = \bigcup_{0 < \alpha < 1} \bigcup_{\omega \in S^1} \{y \in \Omega \mid \lim_{\tau \rightarrow \infty} I_{(y, \omega)}^\alpha(\tau) = 0\},$$

が成り立つ。

定理 4.1 の 2 番目の、すなわち、錐  $C_y(\omega, \pi\alpha/2)$  の閉包と  $D$  の閉包の共通部分が空でなくかつ同時に両方の境界点になっている点のみからなる場合の解析が最も微妙である。これをカバーできているのが良い点であり、したがって系 4.2 の公式が得られるのである。

命題 2.1 の (2.1) によれば、定理 4.1 の証明のためには次の積分の  $\tau \rightarrow \infty$  における対応する性質を確立することが必要十分である：

$$J_{(y, \omega)}(\tau) = \int_D |\nabla e_\tau^\alpha(x; y, \omega)|^2 dx.$$

Mittag-Leffler の関数で直接表現するならば、等式

$$\nabla e_\tau^\alpha(x; y, \omega) = \tau(\omega + i\omega^\perp) E'_\alpha(\tau(x - y) \cdot (\omega + i\omega^\perp))$$

が成り立つから、

$$J_{(y, \omega)}(\tau) = 2\tau^2 \int_D |E'_\alpha(\tau(x - y) \cdot (\omega + i\omega^\perp))|^2 dx$$

を得る。次部分節では Mittag-Leffler の関数の漸近挙動についての知見を述べる。

#### 4.2. Mittag-Leffler の関数の漸近挙動

次の補題が前節の冒頭で述べた Mittag-Leffler の関数の漸近挙動の詳細である。

**補題 4.3.**  $0 < \alpha < 1$ 、 $N \geq 2$  および  $R > 0$  とせよ。

•  $0 < \epsilon < \alpha/2$  とせよ。正の定数  $C_{\alpha, \epsilon, N, R}$  が存在して、もし  $|z| \geq R$  および  $|\arg z| \leq \pi(\alpha/2 + \epsilon)$  ならば各  $j = 0, 1$  に対して評価

$$\left| \frac{d^j}{dz^j} \left\{ E_\alpha(z) - \frac{1}{\alpha} e^{z^{1/\alpha}} + \sum_{m=1}^{N-1} \frac{z^{-m}}{\Gamma(1 - \alpha m)} \right\} \right| \leq \frac{C_{\alpha, \epsilon, N, R}}{|z|^{N+j}} \quad (4.1)$$

が成り立つ。ここに

$$z^{1/\alpha} \equiv |z|^{1/\alpha} e^{i \arg z / \alpha}.$$

•  $0 < \epsilon < \min(\alpha/2, 1 - \alpha)$  とせよ。正の定数  $C_{\alpha, \epsilon, N, R}$  が存在して、もし  $|z| \geq R$  および  $(\alpha/2 + \epsilon)\pi \leq |\arg z| \leq \pi$  ならば各  $j = 0, 1$  に対して評価

$$\left| \frac{d^j}{dz^j} \left\{ E_\alpha(z) + \sum_{m=1}^{N-1} \frac{z^{-m}}{\Gamma(1 - \alpha m)} \right\} \right| \leq \frac{C_{\alpha, \epsilon, N, R}}{|z|^{N+j}} \quad (4.2)$$

が成り立つ。

この証明の出発点は  $1/\Gamma(z)$  の積分表現を与えている Hankel のループ積分である。

**命題 4.4.**  $C$  を複素平面内の  $t = -\infty$  を負の実軸の下方（下岸、 $\arg t = -\pi$ ）から出発して原点を中心とするある円周に沿って反時計回りに回って負の実軸の上方（上岸、 $\arg t = \pi$ ）から再びその出発点に戻る曲線とする。 $|\arg t| \leq \pi$  を満たす  $t$  に対して  $t^{-z} = e^{-z \log t}$  とおく。このとき公式

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i} \int_C e^t t^{-z} dt \quad (4.3)$$

が成り立つ。

証明不等式

$$|t^{-z}| \leq |t|^{-\operatorname{Re} z} e^{\pi |\operatorname{Im} z|}$$

と  $C$  上のすべての点  $t$  は  $|t| \geq A \equiv \operatorname{dis}(C, 0) > 0$  を満たすから、任意に与えられた正数  $K$  に対して正の定数  $C_K$  が存在して  $|z| \leq K$  を満たすすべての  $z$  および  $C$  上のすべての  $t$  に対して

$$|t^{-z}| \leq C_K |t|^K$$

が成り立つ。これは、 $z$  がコンパクト集合を動く限り (4.3) の右辺の積分が一様に絶対収束することを導きしたがって  $z$  の整関数を定義する。 $1/\Gamma(z)$  も整関数であるから  $0 < \operatorname{Re} z < 1$  を満たす  $z$  について (4.3) を証明すればよい。

$\epsilon$  を任意の正の数とする。Cauchy の積分定理によれば、(4.3) の右辺の積分路  $C$  を区間  $]-\infty, -\epsilon]$  の上岸と下岸および円周  $|t| = \epsilon$  からなる曲線  $C_\epsilon$  へ変形できる。 $\epsilon \rightarrow 0$  のときこの円周からの積分への寄与は消えることが評価

$$\left| \int_{|t|=\epsilon} e^t t^{-z} dt \right| \leq 2\pi e^\epsilon \epsilon^{1 - \operatorname{Re} z}$$

より得る。 $\operatorname{Re} z < 1$  が効いていることに注意しよう。したがって  $\epsilon \rightarrow 0$  のとき

$$\begin{aligned} \int_C e^t t^{-z} dt &= \int_{C_\epsilon} e^t t^{-z} dt \\ &\rightarrow - \int_{-\infty}^0 e^{-s} s^{-z} e^{i\pi z} ds - \int_0^{\infty} e^{-s} s^{-z} e^{-i\pi z} ds \\ &= 2i \sin(\pi z) \Gamma(1 - z). \end{aligned} \quad (4.4)$$

を得る。ここで次の表示公式を使用したことに注意する：

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt, \quad \operatorname{Re} z > 0.$$

最後に (4.4) と相反公式

$$\Gamma(z) \Gamma(1 - z) = \frac{\pi}{\sin \pi z}, \quad z \neq 0, \pm 1, \pm 2, \dots$$

から (4.3) が成り立つことが結論される。

□

(4.3) から Mittag-Leffler の積分表示を得よう。与えられた  $\eta \in ]0, \pi/2[$  および  $r > 0$  に対して無限遠点から半直線  $t = \rho e^{-i(\pi/2+\eta)}$ 、 $\rho \geq r$  上原点へ向けて出発し円周の一部  $t = re^{i\theta}$ 、 $|\theta| \leq \pi/2 + \eta$  上を反時計回りに動き次に半直線  $t = \rho e^{i(\pi/2+\eta)}$ 、 $\rho \geq r$  上無限遠点へ到る曲線を  $c(\eta, r)$  であらわす。このとき (4.3) の積分路  $C$  を  $c(\eta, r)$  に変形できることは容易にわかり公式

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i} \int_{c(\eta, r)} e^t t^{-z} dt \quad (4.5)$$

を得る。

次の公式は Mittag-Leffler の積分表示と呼ばれる。

**命題 4.5.** 任意に与えられた  $\eta \in ]0, \pi/2[$  および正数  $r$  に対して公式

$$E_\alpha(z) = \frac{1}{2\pi i} \int_{c(\eta, r)} \frac{t^{\alpha-1} e^t}{t^\alpha - z} dt, \quad |z|^{1/\alpha} < r \quad (4.6)$$

が成り立つ。

**証明** 曲線  $c(\eta, r)$  上のすべての点  $t$  に対して  $|z| < |t^\alpha|$  が成り立つから等式

$$\begin{aligned} \frac{t^{\alpha-1} e^t}{t^\alpha - z} &= \frac{t^{\alpha-1} e^t}{t^\alpha (1 - \frac{z}{t^\alpha})} \\ &= \frac{e^t}{t} \sum_{m=0}^{\infty} \left(\frac{z}{t^\alpha}\right)^m \\ &= \sum_{m=0}^{\infty} \frac{z^m e^t}{t^{\alpha m+1}} \end{aligned}$$

を得る。右辺項別積分し、 $z = \alpha m + 1$  に対する (4.5) を使えば (4.6) を得る。

□

補題 4.3 の証明を与えよう。 $j = 0$  の場合のみ与える。

**(4.1) の証明** 複素数  $z$  は  $|\arg z| \leq \pi(\alpha/2 + \epsilon)$  および  $|z| \geq R$  を満たすとせよ。このとき方程式  $t^\alpha = z$  は  $|\arg t| \leq \pi$  を満たす唯一つの解  $z^{1/\alpha}$  を持つ。詳しくは  $|\arg z^{1/\alpha}| \leq \pi(1/2 + \epsilon/\alpha)$  が成り立つ。この右辺が  $\pi$  より小さいということに  $\epsilon$  の制限  $\epsilon < \alpha/2$  が効いていることに注意しよう。 $z^{1/\alpha}$  は  $t^\alpha - z$  の一位の零点であることに注意すると

$$\lim_{\epsilon \rightarrow 0} \frac{1}{2\pi i} \int_{|t-z^{1/\alpha}|=\epsilon} \frac{t^{\alpha-1} e^t}{t^\alpha - z} dt = \frac{1}{\alpha} e^{z^{1/\alpha}} \quad (4.7)$$

を得る。

$\epsilon'$  を  $\epsilon/\alpha < \epsilon' < 1/2$  を満たすものとし積分路  $c' = c(\pi\epsilon', (R/2)^{1/\alpha})$  を考える。 $z$  に対して  $|z|^{1/\alpha} < r$  を満たす  $r > 0$  をとり、 $\eta = \epsilon\pi/\alpha$  に対する (4.6) から出発してその積分路を  $c'$  へ変形する。 $c(\eta, r)$  と  $c'$  は被積分関数の特異点  $z^{1/\alpha}$  を囲んでいるからコーシーの積分定理を使うとその周りの積分 (4.7) が出てくる。他の寄与は消えることが簡単に確認できて等式

$$E_\alpha(z) = \frac{1}{\alpha} e^{z^{1/\alpha}} + \frac{1}{2\pi i} \int_{c'} \frac{t^{\alpha-1} e^t}{t^\alpha - z} dt \quad (4.8)$$



を得る。

$c'$  上のすべての点  $t$  に対して  $t^\alpha$  は  $z$  と異なるから、等式

$$\frac{t^\alpha}{t^\alpha - z} = - \sum_{m=1}^{N-1} \left(\frac{t^\alpha}{z}\right)^m - (1 - \frac{t^\alpha}{z})^{-1} \left(\frac{t^\alpha}{z}\right)^N, \quad N \geq 2 \quad (4.9)$$

を得る。したがって  $\eta = \pi\epsilon'$  および  $r = (R/2)^{1/\alpha}$  に対する (4.5)、(4.8) および (4.9) から等式

$$E_\alpha(z) = \frac{1}{\alpha} e^{z^{1/\alpha}} - \sum_{m=1}^{N-1} \frac{z^{-m}}{\Gamma(1 - \alpha m)} - \frac{1}{2\pi i} \left\{ \int_{c'} (1 - \frac{t^\alpha}{z})^{-1} t^{\alpha N-1} e^t dt \right\} z^{-N} \quad (4.10)$$

を得る。 $\theta = \arg(t^\alpha/z)$  とおく。

次の二つの主張は丁寧に見ればわかる：

- もし  $|t| \geq (R/2)^{1/\alpha}$  および  $t$  が  $c'$  上にあるならば、 $\theta$  は  $(\alpha\epsilon' - \epsilon)\pi \leq |\theta| \leq \pi(3\alpha/2 + \epsilon)$  を満たす；

- もし  $|t| = (R/2)^{1/\alpha}$  ならば、そのとき  $|t^\alpha/z| \leq 1/2$ 。  
 $0 < (\alpha\epsilon' - \epsilon)\pi$  および  $\pi(3\alpha/2 + \epsilon) < 2\pi\alpha < 2\pi$  を見れば、 $z$  に無関係な正の定数  $C_{\alpha,\epsilon}$  が存在して

$$\inf_{t \in c'} \left| 1 - \frac{t^\alpha}{z} \right| \geq C_{\alpha,\epsilon} \quad (4.11)$$

が成り立つことが結論される。以上の議論で  $\epsilon$  および  $\alpha$  に対する制限  $0 < \epsilon < \alpha/2$  および  $0 < \alpha < 1$  が重要であることが見てとれるであろう。(4.10) および (4.11) から望んだ評価が導かれることは言うまでもない。

□

**(4.2) の証明**  $z$  は  $\pi(\alpha/2 + \epsilon) \leq |\arg z| \leq \pi$  および  $|z| \geq R$  を満たすとせよ。 $c'' = c(\pi\epsilon/(2\alpha), (R/2)^{1/\alpha})$  および  $c''$  上の点  $t$  に対して  $\theta = \arg(t^\alpha/z)$  とおきその動く範囲を調べる。最初に、 $t$  が  $|t| \geq (R/2)^{1/\alpha}$  および  $\arg t = \pi/2(1 + \epsilon/\alpha)$  を満たすとき  $\theta$  の動く範囲を調べると以下のことがわかる：

- もし  $\pi(\alpha/2 + \epsilon) \leq \arg z \leq \pi$  ならば、そのとき  $|\theta + \pi/2| \leq \pi(1 - \alpha - \epsilon)/2$ ;
- もし  $-\pi \leq \arg z \leq -\pi(\alpha/2 + \epsilon)$  ならば、そのとき  $\pi(\alpha + 3\epsilon/2) \leq \theta \leq \pi(1 + \alpha/2 + \epsilon/2)$ 。

次に、 $t$  が  $|t| \geq (R/2)^{1/\alpha}$  および  $\arg t = -\pi/2(1 + \epsilon/\alpha)$  を満たすときを考えると以下のことがわかる：

- もし  $\pi(\alpha/2 + \epsilon) \leq \arg z \leq \pi$  ならば、そのとき  $-\pi(1 + \alpha/2 + \epsilon/2) \leq \theta \leq -\pi(\alpha + 3\epsilon/2)$ ;
- もし  $-\pi \leq \arg z \leq -\pi(\alpha/2 + \epsilon)$  ならば、そのとき  $\pi\epsilon/2 \leq \theta \leq \pi(1 - \alpha/2 - \epsilon/2)$ 。

さらに  $|t| = (R/2)^{1/\alpha}$  ならば  $|t^\alpha/z| < 1/2$  が成り立つ。

これらの観察から、 $z$  に無関係な正の定数  $C'_{\alpha,\epsilon}$  が存在して

$$\inf_{t \in c''} \left| 1 - \frac{t^\alpha}{z} \right| \geq C'_{\alpha,\epsilon} \quad (4.12)$$

が成り立つことがわかる。

さて  $z$  に対して  $r > |z|^{1/\alpha}$  を満たす  $r$  をとる。 $\eta = \pi\epsilon/(2\alpha)$  に対する (4.6) を考える。上の議論から (4.6) の被積分関数の特異点は  $c''$  およびその右側に現れない。したがって積分路を  $c(\eta, r)$  から  $c''$  へ変形できる。そのとき  $c''$  上の各点  $t$  で (4.9) が成り立つから、 $r = (R/2)^{1/\alpha}$  に対する (4.5) から

$$E_\alpha(z) = - \sum_{m=1}^{N-1} \frac{z^{-m}}{\Gamma(1 - \alpha m)} - \frac{1}{2\pi i} \left\{ \int_{c''} (1 - \frac{t^\alpha}{z})^{-1} t^{\alpha N-1} e^t dt \right\} z^{-N}$$

を得る。これと (4.12) から望んだ評価を得る。

□

#### 4.3. 不連続性の可視部の抽出

$\tau \rightarrow \infty$  のときの  $J_{(y, \omega)}(\tau)$  の漸近挙動を三つの場合に分けて考察する。

##### 1. $\overline{\{C_y(\omega, \pi\alpha/2)\}} \cap \overline{D} = \emptyset$ のとき

これは  $\overline{D}$  が  $\mathbf{R}^2 \setminus \overline{C_y(\omega, \pi\alpha/2)}$  に含まれることを意味する。このとき  $\min(\alpha/2, 1-\alpha)$  より小さい正数  $\epsilon$  および正数  $R$  で次の包含関係が成り立つようなものがとれる：

$$\overline{D} \subset \{x \in \mathbf{R}^2 \mid |x - y| \geq R\} \cap \{\mathbf{R}^2 \setminus \overline{C_y(\omega, (\alpha/2 + \epsilon)\pi)}\}.$$

$z = \tau(x - y) \cdot (\omega + i\omega^\perp)$  とおく。このときもし  $x$  が  $\{x \in \mathbf{R}^2 \mid |x - y| \geq R\} \cap \{\mathbf{R}^2 \setminus \overline{C_y(\omega, (\alpha/2 + \epsilon)\pi)}\}$  の要素であれば、 $z$  は  $|z| \geq R\tau$  および  $(\alpha/2 + \epsilon)\pi \leq |\arg z| \leq \pi$  を満たす。したがって、(4.2) が適用できて、十分大きい正数  $\tau_0$  および正の定数  $C$  をとると  $\tau_0$  以上のすべての  $\tau$  に対して評価  $|E'_\alpha(z)| \leq C\tau^{-2}$  を得る。これは  $J(\tau) = O(\tau^{-2})$  をもたらす。

##### 2. $\overline{\{C_y(\omega, \pi\alpha/2)\}} \cap \overline{D} \neq \emptyset$ および $C_y(\omega, \pi\alpha/2) \cap D = \emptyset$ のとき

$y$  を中心とする極座標  $(r, \theta)$  を次のように導入する：

$$x = y + r(\cos \theta \omega + \sin \theta \omega^\perp).$$

このとき等式

$$|\nabla e_\tau^\alpha(x; y, \omega)|^2 = 2\tau^2 |E'_\alpha(\tau r e^{i\theta})|^2$$

が成り立つ。

- 最初に  $y$  が  $\partial D$  上にあるときについて考える。

$\partial D$  の Lipschitz 連続性から、適当な正数  $\delta$  および区間  $[\pi\alpha/2, \pi[ \cup ]-\pi, -\pi\alpha/2[$  の適当な部分区間  $[a, b]$  をとると包含関係

$$C \equiv \{y + r(\cos \theta \omega + \sin \theta \omega^\perp) \mid \theta \in [a, b], 0 < r \leq \delta\} \subset D$$

が成り立つ。したがって評価

$$\begin{aligned} J_{(y, \omega)}(\tau) &\geq \int_C |\nabla e_\tau^\alpha(x; y, \omega)|^2 dx = 2\tau^2 \int_a^b d\theta \int_0^\delta r dr |E'_\alpha(\tau r e^{i\theta})|^2 \\ &= 2 \int_a^b d\theta \int_0^{\tau\delta} |E'_\alpha(s e^{i\theta})|^2 s ds \end{aligned} \quad (4.13)$$

を得る。(4.2) によれば、 $\theta \in [a, b]$  に関して一様に  $s \rightarrow \infty$  のとき  $|E'_\alpha(s e^{i\theta})|^2 = O(s^{-4})$  が成り立つ。これは

$$0 < 2 \int_a^b d\theta \int_0^\infty |E'_\alpha(s e^{i\theta})|^2 s ds < \infty$$

を導く。これと (4.13) より、十分大きい正数  $\tau_0$  に対して  $\inf_{\tau > \tau_0} J_{(y, \omega)}(\tau) > 0$  が成り立つことが結論される。

- 次に  $y$  が  $\partial D$  上にないときについて考える。

この場合が最も繊細な取り扱いを必要とする。 $\overline{C_y(\omega, \pi\alpha/2)} \cap \partial D$  上の点を一つとりそれを  $x_0$  であらわす。 $x_0$  と  $y$  は異なるから、 $x_0$  は

$$x_0 = y + r_0 \left( \cos \frac{\pi\alpha}{2} \omega + \sin \frac{\pi\alpha}{2} \omega^\perp \right)$$

または

$$x_0 = y + r_0 \left( \cos\left(-\frac{\pi\alpha}{2}\right) \omega + \sin\left(-\frac{\pi\alpha}{2}\right) \omega^\perp \right),$$

と書かれるとしてよい。ここで  $r_0 = |x_0 - y|$ 。

後者の場合の取り扱いとは前者のそれと同じであるから前者のみについて述べる。  
 $\partial D$  が Lipschitz 連続であるから、 $\min(\alpha/2, (1-\alpha)/2)$  より小さい正数  $\epsilon$  および  $D$  に含まれる、 $x_0$  を頂点の一つとし、他の 2 頂点が

$$x_1 = y + r_1 \left( \cos\left(\frac{\pi\alpha}{2} + \pi\epsilon\right) \omega + \sin\left(\frac{\pi\alpha}{2} + \pi\epsilon\right) \omega^\perp \right),$$

$$x_2 = y + r_2 \left( \cos\left(\frac{\pi\alpha}{2} + \pi\epsilon\right) \omega + \sin\left(\frac{\pi\alpha}{2} + \pi\epsilon\right) \omega^\perp \right),$$

$$r_2 > r_1$$

で与えられる三角形 (の内部)  $\Delta$  がとれる。

各  $\eta \in [0, \pi\epsilon]$  に対して

$$r_1(\eta) = \frac{r_1 r_0 \sin \pi\epsilon}{r_0 \sin \eta + r_1 \sin(\pi\epsilon - \eta)}, \quad r_2(\eta) = \frac{r_2 r_0 \sin \pi\epsilon}{r_0 \sin \eta + r_2 \sin(\pi\epsilon - \eta)} \quad (4.14)$$

とおく。このとき  $\Delta$  は次のように表示される：

$$\Delta = \{y + r \left( \cos\left(\frac{\pi\alpha}{2} + \eta\right) \omega + \sin\left(\frac{\pi\alpha}{2} + \eta\right) \omega^\perp \right) \mid \eta \in ]0, \pi\epsilon[, r_1(\eta) < r < r_2(\eta)\}.$$

これを使えば評価

$$J_{(y, \omega)}(\tau) \geq 2\tau^2 \int_0^{\pi\epsilon} d\eta \int_{r_1(\eta)}^{r_2(\eta)} |E'_\alpha(\tau r e^{i(\pi\alpha/2 + \eta)})|^2 r dr \quad (4.15)$$

を得る。

(4.1) によれば、評価

$$|E'_\alpha(s e^{i(\pi\alpha/2 + \eta)})|^2 s \geq \frac{C}{2} \{e^{-2s^{1/\alpha} \sin(\eta/\alpha)} s^{2/\alpha-1} - s^{-3}\}, \quad s \geq s_0 \quad (4.16)$$

を得る。ここに  $s_0 \gg 1$  および正の定数  $C$  は  $\eta$  に無関係である。

今  $\tau$  は  $\tau \min(r_1, r_0) \geq s_0$  を満たすとせよ。このときすべての  $r \in ]r_1(\eta), r_2(\eta)[$  に対して  $\tau r \geq s_0$  が成り立つから、(4.15) および (4.16) から評価

$$\begin{aligned} J_{(y, \omega)}(\tau) &\geq C\tau \int_0^{\pi\epsilon} d\eta \int_{r_1(\eta)}^{r_2(\eta)} \{e^{-2\tau^{1/\alpha} r^{1/\alpha} \sin(\eta/\alpha)} \tau^{2/\alpha-1} r^{2/\alpha-1} - \tau^{-3} r^{-3}\} dr \\ &= C\{\tau^{2/\alpha} \int_0^{\pi\epsilon} d\eta \int_{r_1(\eta)}^{r_2(\eta)} e^{-2\tau^{1/\alpha} r^{1/\alpha} \sin(\eta/\alpha)} r^{2/\alpha-1} dr - \tau^{-2} \int_0^{\pi\epsilon} d\eta \int_{r_1(\eta)}^{r_2(\eta)} r^{-3} dr\} \\ &\equiv C\{J_1(\tau) - J_2(\tau)\} \end{aligned} \quad (4.17)$$

を得る。

ここで  $c_1 = \min(r_1, r_0)$  および  $c_2 = \max(r_2, r_0)$  とおく。  $J_2(\tau)$  は

$$J_2(\tau) = \tau^{-2} \int_0^{\pi\epsilon} \frac{1}{2} \left( \frac{1}{r_1(\eta)^2} - \frac{1}{r_2(\eta)^2} \right) d\eta \leq \frac{\pi\epsilon\tau^{-2}}{2} \left\{ \frac{1}{c_1^2} - \frac{1}{c_2^2} \right\} = O(\tau^{-2}) \quad (4.18)$$

と評価される。(4.14) から等式

$$r_2(\eta) - r_1(\eta) = \frac{r_0^2(r_2 - r_1) \sin \pi\epsilon \sin \eta}{\{r_0 \sin \eta + r_2 \sin(\pi\epsilon - \eta)\} \{r_0 \sin \eta + r_1 \sin(\pi\epsilon - \eta)\}}.$$

を得る。

これは  $C' = (r_2 - r_1)r_0^2 \sin \pi\epsilon / 2r_1r_2$  および  $0 < \eta < \pi\epsilon' \ll \pi\epsilon$  を満たす  $\eta$  に対して評価  $r_2(\eta) - r_1(\eta) \geq C'\eta$  をもたらす。そのとき次の評価を得る：

$$\begin{aligned} J_1(\tau) &\geq \tau^{2/\alpha} \int_0^{\pi\epsilon'} d\eta \int_{r_1(\eta)}^{r_2(\eta)} e^{-2\tau^{1/\alpha} c_2^{1/\alpha} \sin(\eta/\alpha)} c_1^{2/\alpha-1} dr \\ &= c_1^{2/\alpha-1} \tau^{2/\alpha} \int_0^{\pi\epsilon'} (r_2(\eta) - r_1(\eta)) e^{-2\tau^{1/\alpha} c_2^{1/\alpha} \sin(\eta/\alpha)} d\eta \\ &\geq C' c_1^{2/\alpha-1} \tau^{2/\alpha} \int_0^{\pi\epsilon'} \eta e^{-2\tau^{1/\alpha} c_2^{1/\alpha} \sin(\eta/\alpha)} d\eta \\ &\geq C' c_1^{2/\alpha-1} \tau^{2/\alpha} \int_0^{\pi\epsilon'} \eta e^{-2\tau^{1/\alpha} c_2^{1/\alpha} \eta/\alpha} d\eta. \end{aligned} \quad (4.19)$$

次の漸近式は容易に得られる：

$$\tau^{2/\alpha} \int_0^{\pi\epsilon'} \eta e^{-2\tau^{1/\alpha} c_2^{1/\alpha} \eta/\alpha} d\eta = (2c_2^{1/\alpha}/\alpha)^{-2} + O(\tau^{1/\alpha} e^{-2\tau^{1/\alpha} c_2^{1/\alpha} \epsilon' \pi/\alpha}). \quad (4.20)$$

(4.19) および (4.20) の組み合わせは  $\liminf_{\tau \rightarrow \infty} J_1(\tau) > 0$  を導く。この事実と (4.17) および (4.18) から望んだ結論を得る。

### 3. $C_y(\omega, \pi\alpha/2) \cap D \neq \emptyset$ のとき

区間  $] -\pi\alpha/2, \pi\alpha/2[$  の部分区間  $[a, b]$  および正数  $R_1, R_2$  で包含関係

$$C' \equiv \{y + r(\cos \theta \omega + \sin \theta \omega^\perp) \mid \theta \in [a, b], R_1 \leq r \leq R_2\} \subset D$$

が成り立つものがとれる。

$c = \min(\cos a, \cos b)(> 0)$  とおくと、(4.1) から評価

$$|E'_\alpha(se^{i\theta})|^2 \geq \frac{1}{2} \left( \frac{s^{1/\alpha-1}}{\alpha^2} e^{s^{1/\alpha}c} \right)^2 - \frac{C}{s^4}, \quad s \geq s_0 \gg 1, \quad \theta \in [a, b]$$

を得る。これは評価

$$\begin{aligned} J_{(y,\omega)}(\tau) &\geq \int_{C'} |\nabla e_\tau^\alpha(x; y, \omega)|^2 dx = 2\tau^2 \int_a^b d\theta \int_{R_1}^{R_2} |E'_\alpha(\tau r e^{i\theta})|^2 r dr \\ &= 2 \int_a^b d\theta \int_{\tau R_1}^{\tau R_2} |E'_\alpha(se^{i\theta})|^2 s ds \\ &\geq \frac{b-a}{\alpha^4} \int_{\tau R_1}^{\tau R_2} s^{2/\alpha-1} e^{2s^{1/\alpha}c} ds - \frac{2C\eta}{\tau^2} \left( \frac{1}{R_1^2} - \frac{1}{R_2^2} \right) \end{aligned}$$

を導く。今  $\tau$  を  $\tau R_1 \geq 1$  となるようにとる。そのとき評価

$$\begin{aligned} (b-a) \int_{\tau R_1}^{\tau R_2} s^{2/\alpha-1} e^{2s^{1/\alpha}c} ds &\geq (b-a) \int_{\tau R_1}^{\tau R_2} s^{1/\alpha-1} e^{2s^{1/\alpha}c} ds \\ &= \frac{(b-a)\alpha}{2c} \int_{\tau R_1}^{\tau R_2} \frac{d}{ds} e^{2s^{1/\alpha}c} ds \\ &= \frac{(b-a)\alpha}{2c} e^{2(\tau R_2)^{1/\alpha}c} \{1 - e^{-2\tau^{1/\alpha}(R_2^{1/\alpha}-R_1^{1/\alpha})c}\} \end{aligned}$$

を得る。これらから評価

$$J_{(y,\omega)}(\tau) \geq C' e^{2(R_2\tau)^{1/\alpha}c}, \quad \tau > \tau_0 \gg 1$$

が得られる。

これで定理 4.1 の証明が完了した。

□

定理 4.1 は三次元においてもしかるべく形で成り立つであろうが、しかし実際に確認しているわけではない。細かい計算を実際にやってみるとよいであろう。

定理 4.1 に相当する、空洞ではなく**未知の導電率をもった介在物**の場合における定理は [40] で確立された。詳しくは、導電率  $\sigma$  が

$$\sigma(x) = \begin{cases} 1, & x \in \Omega \setminus D, \\ 1 + h(x), & x \in D \end{cases}$$

という形で与えられていて、 $h(x)$  は、 $1 + h(x)$  が  $D$  で正の下界をもつ**未知**の本質的に有界な関数であって、 $D$  で正の下界をもつかまたは  $-h(x)$  がそうであるという仮定の下で、導電場の方程式 (1.9) の二次元版においても定理 4.1 に類似な定理が成り立つ。[50] では、その定理を基礎にした介在物の位置や形についての情報を抽出するためのアルゴリズムを提唱し、その数値的検証を実行した。その結果、位置や個数はよく再現できるという知見を得た。我々は、この方法は、他のさまざまな間接再構成の方法へ良い初期推測を提供するであろうと期待している。

最後に、数学的に最も興味がある問題を述べよう。その叙述は単純である。**Mittag-Leffler** の関数を用いて**定理 2.3 以上のことがわかるか**。具体的には、 $H^{1/2}(\partial\Omega)$  の、定数関数でない要素  $f$  を任意に与え固定し、 $\tau$  の関数

$$\langle (\Lambda_0 - \Lambda_D)f, e_\tau^\alpha(\cdot; y, \omega)|_{\partial\Omega} \rangle$$

の  $\tau \rightarrow \infty$  のときの挙動を明らかにせよ。

## 5 因数分解法

Kirsch[54] の因数分解法 (**Factorization method**) は、一言でいうと、導体そのものをその中の未知の不連続性ととも、導体の境界上のある関数空間の中に忠実にコピーし、不連続性のコピーのみを観測データの因数分解を通じて抽出する方法である。現在彼の方法はさまざまな逆散乱問題および境界値逆問題に応用されている。この節では彼のアイデアの核心部分を紹介しよう。

### 5.1. 因数分解公式

彼の方法は、私の見たところ、観測データとして Dirichlet-to-Neumann 写像よりは Neumann-to-Dirichlet 写像を取ったほうが自然である。そこでこの節では観測データとして Neumann-to-Dirichlet 写像をとり、まず基礎となる**因数分解公式の作り方**に焦点を合わせて解説する。アイデア中心であるので、細かい関数空間の設定や定義の詳細についての記述は省略するが、それをきちんと記述することは今後の発展にとって有用と思われるので、関心のある方は良い演習問題とおもってそれを実行することをお勧めする。なお以下では  $\nu$  は  $\Omega \setminus \bar{D}$  における単位法線ベクトル場であるが、 $\nu$  は  $\partial\Omega$  上  $\Omega$  に対して外向きであるのは以前と同じであるが、 $\partial D$  上は  $D$  に対して外向きであるようにしてあることを注意しておく。

まず背景となる Neumann-to-Dirichlet 写像  $R_0$  は、 $\Omega$  における調和関数  $v$  でその  $\partial\Omega$  上へのトレースの  $\partial\Omega$  全体での積分が消えるものに対して、

$$R_0\left(\frac{\partial v}{\partial \nu}\Big|_{\partial\Omega}\right) = v|_{\partial\Omega}$$

を満たすものとして定義される。

次に**空洞がある場合**の Neumann-to-Dirichlet 写像  $R_D$  は、 $\Omega \setminus \bar{D}$  における調和関数  $u$  でその  $\partial D$  における法線微分が消え、さらにその  $\partial\Omega$  上へのトレースの  $\partial\Omega$  全体での積分が消えるものに対して、

$$R_D\left(\frac{\partial u}{\partial \nu}\Big|_{\partial\Omega}\right) = u|_{\partial\Omega}$$

を満たすものとして定義される。

さて任意に与えられた、 $\partial\Omega$  全体での積分が消える関数  $g$  に対して、上の  $u$ 、 $v$  として、

$$g = \frac{\partial u}{\partial \nu}\Big|_{\partial\Omega} = \frac{\partial v}{\partial \nu}\Big|_{\partial\Omega}$$

を満たすものとする。 $f$  は任意に与えられた、 $\partial\Omega$  全体での積分が消える関数とする。この  $f$  に対して、 $\Omega$  における調和関数  $v'$  でその  $\partial\Omega$  上へのトレースの  $\partial\Omega$  全体での積分が消えかつ

$$f = \frac{\partial v'}{\partial \nu}\Big|_{\partial\Omega}$$

を満たすものとする。このとき部分積分によれば等式

$$\int_{\partial\Omega} f(R_D - R_0)g dS = \int_{\partial D} \frac{\partial v'}{\partial \nu} u dS \quad (5.1)$$

を得る。これは命題 2.4 の (2.10) に対応していることを注意する。

ここで  $D$  における調和関数でその  $\partial D$  上へのトレースの  $\partial D$  全体における積分が消えるものの  $\partial D$  における Neumann データを同じく  $\partial D$  上の Dirichlet データへ対応させる作用素を  $\tilde{\Lambda}_-$  とあらわそう。このとき等式

$$v|_{\partial D} - \frac{\int_{\partial D} v dS}{|\partial D|} = \tilde{\Lambda}_-\left(\frac{\partial v}{\partial \nu}\Big|_{\partial D}\right) \quad (5.2)$$

が成り立つ。

次に  $\Omega$  における調和関数の  $\partial\Omega$  における Neumann データを  $\partial D$  における Neumann データに対応させる作用素を  $H_1$  であらわそう。したがって等式

$$\frac{\partial v'}{\partial \nu}|_{\partial D} = H_1 f \quad (5.3)$$

および

$$\frac{\partial v}{\partial \nu}|_{\partial D} = H_1 g \quad (5.4)$$

が成り立つことになる。(5.2), (5.3) および (5.4) より等式

$$\int_{\partial D} \frac{\partial v'}{\partial \nu} v dS = \int_{\partial D} H_1 f \tilde{\Lambda}_- H_1 g dS \quad (5.5)$$

を得る。ここで  $v'$  は  $D$  で調和であることを使っていることに注意しよう。

次に  $w = u - v$  とおこう。 $\Omega \setminus \bar{D}$  における調和関数で、その  $\partial\Omega$  における Neumann データが消えかつその  $\partial D$  上へのトレースの  $\partial D$  全体の積分が消えるものに対して、その  $\partial D$  上の Neumann データを同じく  $\partial D$  上の Dirichlet データへ対応させる作用素を  $\tilde{\Lambda}_+$  であらわそう。等式

$$w|_{\partial D} - \frac{\int_{\partial D} w dS}{|\partial D|} = \tilde{\Lambda}_+ \left( \frac{\partial w}{\partial \nu} \Big|_{\partial D} \right) \quad (5.6)$$

が成り立つことは言うまでもない。 $w$  の  $\partial D$  における Neumann データは

$$\frac{\partial w}{\partial \nu} \Big|_{\partial D} = - \frac{\partial v}{\partial \nu} \Big|_{\partial D}$$

であった。これと (5.3)、(5.4) および (5.6) より等式

$$\int_{\partial D} \frac{\partial v'}{\partial \nu} w dS = - \int_{\partial D} H_1 f \tilde{\Lambda}_+ H_1 g dS \quad (5.7)$$

を得る。ここでも  $v'$  は  $D$  で調和であることを使っていることに注意しよう。(5.1)、(5.5) および (5.7) から  $R_D - R_0$  の因数分解公式の一手手前の等式

$$\int_{\partial\Omega} f(R_D - R_0) g dS = \int_{\partial D} H_1 f (\tilde{\Lambda}_- - \tilde{\Lambda}_+) H_1 g dS \quad (5.8)$$

を手に入れた。ここまでは、すんなりと来たのであるがこの後、 $\partial D$  上の積分が消えている  $\partial D$  上の任意の関数  $h$  に対して、 $\partial\Omega$  上の積分が消える  $\partial\Omega$  上の関数を対応させる作用素  $H_2$  で等式

$$\int_{\partial D} (H_1 f) h dS = \int_{\partial\Omega} f (H_2 h) dS \quad (5.9)$$

を満たすものを求める必要がある。これができれば (5.8) から第一の**因数分解公式**

$$R_D - R_0 = H_2 (\tilde{\Lambda}_- - \tilde{\Lambda}_+) H_1 \quad (5.10)$$

を得たことになる。**天下りの記述ではない**、 $H_2$  の求め方のこつについていろいろ悩んでいたのだが、ようやく掴んだようであるので、それをここで紹介しよう。要は積分

$$\int_{\partial D} \frac{\partial v'}{\partial \nu} \cdots dS$$

を

$$\int_{\partial\Omega} \frac{\partial v'}{\partial \nu} \cdots dS$$

に書き換えることが出来ればよいと腹をくくる。とすると、 $\Omega \setminus \overline{D}$  における関数をからめて部分積分にもっていけばよいであろうと思いつく。今  $h_+$  を  $\Omega \setminus \overline{D}$  における調和関数としよう。部分積分によれば、等式

$$\int_{\partial D} \frac{\partial v'}{\partial \nu} h_+ dS = \int_{\partial\Omega} \frac{\partial v'}{\partial \nu} h_+ dS - \int_{\Omega \setminus \overline{D}} \nabla v' \cdot \nabla h_+ dx \quad (5.11)$$

を得る。この第二項が邪魔である。とりあえずこれを  $\partial D$  上の積分に書き換えるため、 $h_+$  に境界条件

$$\frac{\partial h_+}{\partial \nu} \Big|_{\partial\Omega} = 0 \quad (5.12)$$

を課そう。すると再び部分積分により等式

$$\int_{\Omega \setminus \overline{D}} \nabla v' \cdot \nabla h_+ dx = - \int_{\partial D} v' \frac{\partial h_+}{\partial \nu} dS \quad (5.13)$$

を得る。しかしこの積分には  $v'$  の法線微分はあらわれない。そこで  $D$  における調和関数  $h_-$  を考える。このとき等式

$$\int_{\partial D} v' \frac{\partial h_-}{\partial \nu} dS = \int_{\partial D} \frac{\partial v'}{\partial \nu} h_- dS \quad (5.14)$$

が成り立つのは言うまでもない。そこで  $h_-$  に境界条件

$$\frac{\partial h_-}{\partial \nu} \Big|_{\partial D} = \frac{\partial h_+}{\partial \nu} \Big|_{\partial D} \quad (5.15)$$

課せば、(5.14) から (5.13) の右辺は  $v'$  の法線微分を含んだ積分になり、(5.11) より等式

$$\int_{\partial D} \frac{\partial v'}{\partial \nu} (h_+ - h_-) dS = \int_{\partial\Omega} \frac{\partial v'}{\partial \nu} h_+ dS \quad (5.16)$$

を得たことになる。そこで与えられた  $h$  に対して  $\Omega \setminus \overline{D}$ 、 $D$  それぞれにおける調和関数  $h_+$ 、 $h_-$  で境界条件 (5.12), (5.15) および

$$h_+ - h_- = h \text{ on } \partial D \quad (5.17)$$

を満たすものを取ってくれば、(5.16) より目的の等式

$$\int_{\partial D} \frac{\partial v'}{\partial \nu} h dS = \int_{\partial\Omega} \frac{\partial v'}{\partial \nu} h_+ dS$$

を得たことになる。 $h_+$  および  $h_-$  それぞれに同じ定数を付け加えても境界条件 (5.12)、(5.15) および (5.17) は満たされるから、 $h_+$  の  $\partial\Omega$  上へのトレースの  $\partial\Omega$  全体での積分は消えているようにできる。そこで作用素  $H_2$  を

$$H_2 h = h_+ \Big|_{\partial\Omega}$$



と定義すれば (5.9) が満たされることがわかる。問題は、 $h_+$ 、 $h_-$  の存在と一意性についてであるが、これは古典的で、有名な本 [61] にゆだねておく。

因数分解公式のもうひとつである  $H_2$  の因数分解公式が重要な役割を演じるのでそれを述べよう。 $\Omega \setminus \overline{D}$  における調和関数で、その  $\partial\Omega$  における法線微分が消えかつその  $\partial\Omega$  上へのトレースの  $\partial\Omega$  における積分が消えるものの  $\partial D$  における Neumann データを  $\partial\Omega$  における Dirichlet データに対応させる作用素を  $G$  であらわそう。 $h_+$  の性質を思い出せば、

$$h_+|_{\partial\Omega} = G\left(\frac{\partial h_+}{\partial\nu}\Big|_{\partial D}\right)$$

が成り立つ。作用素

$$Kh = \frac{\partial h_+}{\partial\nu}\Big|_{\partial D} \quad (5.18)$$

を導入すると

$$H_2 = GK$$

と書かれることになる。鍵は  $K$  が**全単射**であることで、これが言えれば (はっきり定義域および値域を述べていないが) 等式

$$\text{Rang } H_2 = \text{Range } G \quad (5.19)$$

が従うことになる。

まず  $K$  が**単射**であることを示そう。 $Kh = 0$  を仮定しよう。(5.18) より  $h_+$  は  $\Omega \setminus \overline{D}$  全体で定数となるが  $h_+$  の  $\partial\Omega$  上の積分が消えることから  $\Omega \setminus \overline{D}$  全体で  $h_+ = 0$  となる。一方 (5.15) より  $h_-$  は  $D$  で定数関数となるが、(5.17) を  $\partial D$  上で積分することにより、 $\partial D$  上  $h_- = 0$  となることから、 $h$  の  $\partial D$  上の積分が消えることより結論され、したがって再び (5.17) より  $h = 0$  を得る。

つぎに  $K$  が**全射**であることを示そう。 $\partial D$  全体での積分が消えるような  $\partial D$  上の関数  $\phi$  を任意に与える。次の境界値問題の解を  $\Psi_+$  と書こう：

$$\Delta w = 0 \text{ in } \Omega \setminus \overline{D},$$

$$\frac{\partial w}{\partial\nu} = \phi \text{ on } \partial D,$$

$$\frac{\partial w}{\partial\nu} = 0 \text{ on } \partial\Omega,$$

$$\int_{\partial\Omega} w dS = 0.$$

$D$  における次の境界値問題の解を  $\Psi_-$  と書こう：

$$\Delta w = 0 \text{ in } D,$$

$$\frac{\partial w}{\partial\nu} = \phi \text{ on } \partial D,$$

$$\int_{\partial D} w dS = \int_{\partial D} \Psi^+ dS.$$

このとき  $h = \Psi^+|_{\partial D} - \Psi^-|_{\partial D}$  とおくと、 $h$  の  $\partial D$  全体での積分は消えて、解の一意性から  $(\Psi^+, \Psi^-) = (h_+, h_-)$  であることを得て、 $Kh = \phi$  を結論する。

## 5.2. 因数分解公式の一つの帰結

これから先は形式論ですみそうもないので関数空間を導入する：

$$X(\partial\Omega) = \{f \in H^{1/2}(\partial\Omega) \mid \int_{\partial\Omega} f dS = 0\},$$

$$X(\partial D) = \{g \in H^{1/2}(\partial D) \mid \int_{\partial D} g dS = 0\},$$

$$X(\partial\Omega)' = \{f \in H^{-1/2}(\partial\Omega) \mid \langle f, 1 \rangle = 0\},$$

$$X(\partial D)' = \{g \in H^{-1/2}(\partial D) \mid \langle g, 1 \rangle = 0\}$$

$X(\partial\Omega)'$  および  $X(\partial D)'$  は自然に、 $X(\partial\Omega)$ 、 $X(\partial D)$  それぞれの双対空間と同一視できるとに注意しよう。前節のような議論を丁寧にとまとめると以下のような事実を得る：

- $R_D - R_0 : X(\partial\Omega)' \rightarrow X(\partial\Omega)$  は有界、**単射**、非負そしてその dual([76]) は自分自身と一致する；

- $H_1 : X(\partial\Omega)' \rightarrow X(\partial D)'$  は有界かつ単射；

- $H_2 : X(\partial D) \rightarrow X(\partial\Omega)$  は有界；

- $G : X(\partial D)' \rightarrow X(\partial\Omega)$  は有界かつ単射；

- $K : X(\partial D) \rightarrow X(\partial D)'$  は有界かつ全単射；

- $H_2 = GK$ ；

- $H_1$  は  $H_2$  の dual；

- $\tilde{\Lambda}_- - \tilde{\Lambda}_+ : X(\partial D)' \rightarrow X(\partial D)$  は有界、その dual は自分自身と一致し、正定値すなわち、正定数  $C$  が存在して  $X(\partial D)'$  のすべての要素  $\varphi$  に対して不等式

$$\langle \varphi, (\tilde{\Lambda}_- - \tilde{\Lambda}_+)\varphi \rangle \geq C \|\varphi\|_{X(\partial D)'}^2$$

が成り立つ。

この最後の主張は、 $\tilde{\Lambda}_-$  が正定値であること、そして  $-\tilde{\Lambda}_+$  が非負であることから結論される。

問題は因数分解公式 (5.10) および (5.19) から何が従うかである。[56] では次の定理を基礎におく。

**定理 5.1.**  $X_1, X_2$  を反射的 Banach 空間、 $X_1', X_2'$  でそれぞれの双対空間をあらわそう。有界線形作用素  $A : X_1' \rightarrow X_1$ 、 $B : X_1 \rightarrow X_2$  および  $F : X_2' \rightarrow X_2$  は次の三つの条件を満たすとせよ：

- 等式

$$F = BA'B'$$

がなりたつ。ここで  $A', B'$  は  $A, B$  それぞれの dual である；

- $B$  は単射である；

- 正定数  $C$  が存在して  $X_1'$  のすべての要素  $\varphi$  に対して不等式

$$|\langle \varphi, A\varphi \rangle| \geq C \|\varphi\|_{X_1'}^2$$

が成り立つ。

そのとき、 $X_2$  の任意に与えられた要素  $\phi \neq 0$  に対して、 $\phi$  が  $\text{Range } B$  に属するための必要十分条件は  $W(\phi; F)$  が正の値をとることである。ここで  $W : X_2 \rightarrow \mathbb{R}$  は

$$W(\phi; F) = \inf \{ | \langle \Psi, F\Psi \rangle | \mid \Psi \in X_2', \langle \Psi, \phi \rangle = 1 \}$$

と定義される。

注意として、任意に与えられた  $\phi \neq 0$  に対して集合

$$\{ | \langle \Psi, F\Psi \rangle | \mid \Psi \in X_2', \langle \Psi, \phi \rangle = 1 \}$$

が空でないことは、**Hahn-Banach** の拡張定理からの帰結であるが、 $X_2$  が Hilbert 空間のときにはその使用は避けられる。

証明最初に  $\phi$  が  $\text{Range } B$  の要素である場合を考えよう。 $X_1$  に属する 0 でない適当な  $\varphi$  を取ると、 $\phi = B\varphi$  が成り立つ。このとき  $X_2'$  の任意の要素  $\Psi$  に対して不等式

$$\begin{aligned} | \langle \Psi, F\Psi \rangle | &= | \langle \Psi, BA'B'\Psi \rangle | \\ &= | \langle B'\Psi, A'B'\Psi \rangle | \\ &\geq C \|B'\Psi\|_{X_1'}^2 \\ &= C \|B'\Psi\|_{X_1'}^2 \|\varphi\|_{X_1}^2 \|\varphi\|_{X_1}^{-2} \\ &\geq C | \langle B'\Psi, \varphi \rangle |^2 \|\varphi\|_{X_1}^{-2} = C | \langle \Psi, B\varphi \rangle |^2 \|\varphi\|_{X_1}^{-2} \\ &= C | \langle \Psi, \phi \rangle |^2 \|\varphi\|_{X_1}^{-2} \end{aligned}$$

を得る。したがって評価

$$W(\phi; F) \geq C \|\varphi\|_{X_1}^{-2} > 0$$

を得る。

次に  $\phi$  が  $\text{Rang } B$  に属さない場合を考えよう。

$$V = \{ f \in X_2' \mid \langle f, \phi \rangle = 0 \}$$

とおこう。このとき  $B'(V)$  は  $X_1'$  で稠密である。まずこれを証明しよう。 $(X_1')'$  は  $X_1$  と同一視できる故、 $X_1$  の要素  $\varphi$  を任意に与える。 $V$  の任意の要素  $f$  に対して、

$$\langle B'f, \varphi \rangle = 0$$

が成り立つならば、このとき  $\langle f, B\varphi \rangle = 0$  が成り立つ。今  $B\varphi \neq 0$  であったとする。Hahn-Banach の拡張定理によれば、 $B\varphi$  と  $\phi$  が一次独立ではあり得ない。さらに、 $\phi$  は  $\text{Rang } B$  の要素ではないので、 $B\varphi = 0$  でなければならず矛盾である。したがって  $B\varphi = 0$  でなければならず、 $B$  は単射であるから、 $\varphi = 0$  を結論する。これで  $B'(V)$  は  $X_1'$  で稠密であることが分かった。次に、 $X_2'$  の要素  $\phi'$  で

$$\langle \phi', \phi \rangle = 1$$

を満たすものをとる。このとき、 $V$  中の要素からなる列  $\{f_n\}$  で、

$$B' f_n \longrightarrow -B' \phi' \text{ in } X'_1$$

を満たすものが存在することが上で述べたことより結論される。 $X'_2$  中の要素からなる列  $\{\Psi_n\}$  を

$$\Psi_n = f_n + \phi'$$

により定義すると、 $\langle \Psi_n, \phi \rangle = 1$  が成り立つ。さらに

$$W(\phi; F) \leq |\langle \Psi_n, F \Psi_n \rangle|$$

$$= |\langle B' \Psi_n, A' B' \Psi_n \rangle| \leq \|A'\| \|B' \Psi_n\|_{X'_1}^2 \longrightarrow 0$$

を得る。

□

この定理を  $A = \tilde{\Lambda}_- - \tilde{\Lambda}_+$ 、 $B = H_2$  および  $F = R_D - R_0$  に適用する。(5.10)、(5.19) およびその他の考察より、 $\text{Range } G$  の  $R_D - R_0$  のみを使用した特徴付けを得る。

**定理 5.2.**  $X(\partial\Omega)$  の任意に与えられた要素  $\phi \neq 0$  に対して、 $\phi$  が  $\text{Range } G$  に属するための必要十分条件は  $W(\phi; R_D - R_0)$  が正の値をとることである。

ではなぜ  $\text{Rang } G$  の情報が重要であることを説明しなければならない。

$\Omega$  内に任意に与えられた点  $x$  に対して次の境界値問題の解を  $N(\cdot, x)$  と書こう：

$$\Delta_y N(y, x) + \delta(y - x) = 0 \text{ in } \Omega,$$

$$\frac{\partial}{\partial \nu_y} N(y, x) = -\frac{1}{|\partial\Omega|} \text{ on } \partial\Omega$$

$$\int_{\partial\Omega} N(y, x) dS(y) = 0.$$

関数  $N(x, y)$  は **Neumann 関数** と呼ばれ、その構成は

$$\Delta_y E(y, x) = 0 \text{ in } \Omega,$$

$$\frac{\partial}{\partial \nu_y} E(y, x) = -\frac{1}{|\partial\Omega|} - \frac{\partial}{\partial \nu_y} G(y - x) \text{ on } \partial\Omega,$$

$$\int_{\partial\Omega} E(y, x) dS(y) = -\int_{\partial\Omega} G(y - x) dS(y)$$

の一意解  $E(\cdot, x)$  を使って

$$N(y, x) = G(y - x) + E(y, x)$$

という形でなされる。

方向  $\omega \in S^1$  を一つ選び固定する。そして

$$g_{x,\omega}(y) = \omega \cdot \nabla_x N(y, x)|_{\partial\Omega}$$

とおこう。  $g_{x,\omega} \in X(\partial\Omega)$  であり、  $w_x(y) = \omega \cdot \nabla_x N(y, x)$  で定義される  $y \in \Omega$  の関数  $w_x$  は次の Laplace 方程式に対する **Cauchy 問題** の解である：

$$\Delta w = 0 \text{ in } \Omega \setminus \{x\},$$

$$\frac{\partial w}{\partial \nu} = 0 \text{ on } \partial\Omega,$$

$$w = g_{x,\omega} \text{ on } \partial\Omega$$

この関数  $w_x$  は点  $x$  において特異性を持っていることに注意しよう。したがって Cauchy 問題の解の一意性により対応

$$\Omega \ni x \mapsto g_{x,\omega} \in X(\partial\Omega)$$

が一对一であることを得る。これは、 $\Omega$  内の各点が境界上の関数空間の中の要素として**忠実に翻訳された**ことを意味する。そこで空洞  $D$  をこの方法で翻訳したらどうなるか。これに対する答えが次の定理である。

**定理 5.3.** 任意に与えられた  $\Omega$  内の点  $x$  に対して  $x$  が  $D$  の要素であるための必要十分条件は  $g_{x,\omega}$  が  $\text{Rang } G$  に属することである。

**証明**  $x \in D$  のときは、 $w_x$  は  $\Omega \setminus \overline{D}$  で調和で、 $\partial w_x / \partial \nu|_{\partial D} \in X(\partial D)'$  であるから、等式

$$g_{x,\omega} = G\left(\frac{\partial w_x}{\partial \nu}\Big|_{\partial D}\right)$$

を得て  $g_{x,\omega} \in \text{Rang } G$  を得る。 $x \in \Omega \setminus D$  のとき、 $g_{x,\omega} \in \text{Rang } G$  であったとする。このとき  $G$  の定め方によれば、 $\Omega \setminus \overline{D}$  における調和関数で  $w_x$  と同じ  $\partial\Omega$  上の Cauchy データをもつものの存在が結論される。Cauchy 問題の解の一意性を使えば、その調和関数は  $w_x$  と  $x$  以外で一致し、したがって  $x$  で、粗く言って、非有界になる。これは矛盾である。したがって  $g_{x,\omega}$  は  $\text{Rang } G$  に属さない。

□

この定理 (の考え方) は、Colton-Kirsch の Linear sampling method [11] においても知られていたことを注意しよう。要は、 $x$  が  $g_{x,\omega}$  に、 $D$  が  $\text{Rang } G$  に忠実に翻訳されたということである。すると後は観測データ  $R_D - R_0$  から  $\text{Rang } G$  が得られればよいがそれは定理 5.2 そのものである。

定理 5.2 および 5.3 の系として、空洞の観測データによる特徴づけを得る。

**系 5.4.** 任意に与えられた  $\Omega$  内の点  $x$  に対して  $x$  が  $D$  の要素であるための必要十分条件は

$$W(g_{x,\omega}; R_D - R_0) > 0.$$

この系は Kirsch の因数分解法の帰結の一つ ([56]) であり、[54] で導入された初期のもの、すなわち  $|R_D - R_0|^{1/2}$  をしかるべく定義して、その像から  $\text{Rang } G$  を抽出する方法と比較して主張が弱い、その分、議論の構造は単純明快であり、適用可能性が拡大されている。因数分解法の問題への応用については、[4, 5, 6, 21, 15, 23, 55, 57] を参照されたい。とにかく、正直に、素晴らしい方法であると感心するが、一つ違和感をもつとすれば、 $g_{x,\omega}$  にある方向  $\omega$  は何でも良いという意味で、 $\omega$  の役割が死んでいることである。何かこれを生かした方法はあるであろうか。その点系 3.5 あるいは系 4.2 では、不連続性が

ない方向まで込めた情報が抽出されている。このように、因数分解法と探針法あるいは囲い込み法との関係は超関数に対する**特異台**と**波面集合**の関係を想起させる。

## 第二章 体積（面積）のない不連続性を抽出すること

この章では、前章で紹介した探針法と囲い込み法の考え方を、導体内に発生した亀裂や音を伝える媒質内にある薄い障害物などの不連続性の位置およびその形についての情報を観測データから抽出する問題へ適用する。これら不連続性は体積（面積）がないため、第一章と比べてより数学上の技巧を必要とする。

## 6 探針法と亀裂の逆問題

### 6.1. 亀裂の逆問題

最初に問題を定式化しよう。導体  $\Omega$  の中に発生した亀裂を  $\Sigma$  であらわす。 $\Sigma$  は、空でない連結な  $\mathbf{R}^2$  の閉集合で、 $\Omega$  に含まれる一つの Lipschitz 連続な単純閉曲線  $C$  上の異なる二点  $P, Q$  を結ぶ曲線上の弧で与えられると仮定する。後半の仮定は記述の簡単のためであり、この節の方法はもっと複雑な状況へ適用することができることを注意しておく。 $\Sigma$  からその端点を除いた部分を  $\text{Int } \Sigma$ 、端点のみからなる集合を  $\partial\Sigma$  と書く。 $C$  は  $\Omega$  を、 $C$  によって囲まれる領域  $D$  とその外部  $\Omega \setminus \overline{D}$  にわける。

もちろん、 $C$  の  $\Sigma$  以外の部分を変形すれば、上の性質をもつ単純閉曲線は無限に多くあることがわかるであろう。

順問題を記述するため記号および関数空間を導入する。

まず  $\partial D = \Gamma$  とおく。さらに  $\Omega_+ = \Omega \setminus \overline{D}$  および  $D = \Omega_-$  と書こう。

各  $v \in L^2(\Omega)$  に対して  $v_+ = v|_{\Omega_+}$ ,  $v_- = v|_{\Omega_-}$  とおく。

線形空間

$$X(\Omega \setminus \Sigma; D) = \{v \in L^2(\Omega) \mid v_+ \in H^1(\Omega_+), v_- \in H^1(\Omega_-), v_+|_{\Gamma \setminus \Sigma} = v_-|_{\Gamma \setminus \Sigma}\}$$

およびその上のノルムを

$$\|v\|_{X(\Omega \setminus \Sigma; D)} = \|v_+\|_{H^1(\Omega_+)} + \|v_-\|_{H^1(\Omega_-)}$$

により定義する。 $X(\Omega \setminus \Sigma; D)$  はノルム  $\|\cdot\|_{X(\Omega \setminus \Sigma; D)}$  に関して完備である。その閉部分空間を

$$X_0(\Omega \setminus \Sigma; D) = \{v \in X(\Omega \setminus \Sigma; D) \mid v = 0 \text{ on } \partial\Omega\}$$

により定義する。

**定義 6.1.** 任意に与えられた  $f \in H^{1/2}(\partial\Omega)$  に対して、 $u \in X(\Omega \setminus \Sigma; D)$  が境界値問題

$$\begin{aligned} \Delta u &= 0 \text{ in } \Omega \setminus \Sigma, \\ \frac{\partial u}{\partial \nu} &= 0 \text{ on } \Sigma, \\ u &= f \text{ on } \partial\Omega \end{aligned} \tag{6.1}$$

の弱解であるとは、 $u$  の  $\partial\Omega$  上へのトレースが  $f$  と一致し、すべての  $\varphi \in X_0(\Omega \setminus \Sigma; D)$  に対して

$$\int_{\Omega \setminus \Sigma} \nabla u \cdot \nabla \varphi dy = 0 \quad (6.2)$$

が成り立つときを言う。

出発点は (6.1) の弱解の存在と一意性、解が  $D$  の選び方によらないということを確立することである。それを証明することは良い演習問題であり、主張だけまとめておこう。

**命題 6.1.** 任意に固定した  $D$  に対して (5.1) の弱解が唯一つ存在する。さらに解は  $D$  の選び方によらない。

任意に選んだ  $D$  に対して  $H^{1/2}(\partial\Omega)$  上の有界線形汎関数  $\Lambda_\Sigma f$  を公式

$$\langle \Lambda_\Sigma f, h \rangle = \int_{\Omega \setminus \Sigma} \nabla u \cdot \nabla v dy, \quad h \in H^{1/2}(\partial\Omega) \quad (6.3)$$

により定義する。ここに  $u$  は (5.1) の弱解そして  $v$  はその  $\partial\Omega$  上へのトレースが  $h$  と一致する任意の  $X(\Omega \setminus \Sigma; D)$  の要素である。

写像  $\Lambda_\Sigma : f \mapsto \Lambda_\Sigma f$  は **Dirichlet-to-Neumann 写像** と呼ばれる。命題 6.1 および (6.3) によれば  $\Lambda_\Sigma$  も  $D$  の選び方によらないことがわかる。これが後で重要なポイントになっていることをみるであろう。

この節では次の問題を考察する。

**逆問題 6.1.**  $\Lambda_\Sigma$  あるいはその一部から  $\Sigma$  の位置および形についての情報を抽出せよ。

一意性は [17] で証明された。次部分節でこの問題へ探針法を適用する。その前に空洞の場合と比べて困難がどこにあらわれるかを簡単にのべておこう。空洞においては等式 (2.2) があり、それによりすべてが、 $\Omega$  全体で調和な関数  $v$  の空洞上でのエネルギーを調べることに直ちに帰着された。これは空洞が体積（面積）をもっているからであり、亀裂の場合次の等式を得る（演習問題とする）：

$$\langle (\Lambda_0 - \Lambda_D)f, \bar{f} \rangle = \int_{\Omega \setminus \Sigma} |\nabla(u - v)|^2 dx. \quad (6.4)$$

ここで  $u$  は (6.1) の弱解、 $v$  は (1.12) の弱解である。これは (2.2) の空洞が亀裂に退化した場合と考えられる。空洞の場合と違って、今度は  $v$  の寄与があからさまには見えず、直接  $u - v$  のエネルギーについて必要な知見を得なければならない。ここをどうするかが、数学として面白いところである。

## 6.2. 指示列および指示関数

この部分節では探針法を適用して得られる結果を述べよう。

**定義 6.1.** 任意に与えられた  $x \in \Omega$ 、 $x$  を先端にもつ針  $\sigma$  および  $(x, \sigma)$  に対する任意の針列  $\xi = \{v_n\}$  に対して指示列を

$$I(x, \sigma, \xi)_n = \langle (\Lambda_0 - \Lambda_\Sigma)f_n, \bar{f}_n \rangle$$

により定義する。ここで  $f_n$  は  $v_n$  の  $\partial\Omega$  上へのトレースである。

**定義 6.2.** 亀裂の外部  $\Omega \setminus \Sigma$  で定義された関数

$$I(x) = \int_{\Omega \setminus \Sigma} |\nabla w_x|^2 dy, \quad x \in \Omega \setminus \Sigma$$

を指示関数と呼ぶ。ここで  $w_x \in X_0(\Omega \setminus \Sigma; D)$  は次の境界値問題の弱解で  $\Sigma$  による反射解と呼ぶ：

$$\Delta w = 0 \text{ in } \Omega \setminus \Sigma,$$

$$\frac{\partial w}{\partial \nu} = -\frac{\partial}{\partial \nu}(G(\cdot - x)) \text{ on } \Sigma,$$

$$w = 0 \text{ on } \partial\Omega$$

弱解をきちんと定義し、その解の存在と一意性を証明することは演習問題とする。指示関数  $I(x)$  は  $x$  が  $\text{Int } \Sigma$  上の与えられた点に限りなく近づくとき無限大に発散する。他の知見もまとめると次の定理になる：

**定理 6.1.A.**

- (A.1) 任意に与えられた  $\Omega \setminus \Sigma$  内の点  $x$  および  $x$  を先端にもつ針  $\sigma$  に対してもし  $\sigma([0, 1])$  が  $\Sigma$  と共有点を持たないならば、そのとき、 $(x, \sigma)$  に対する任意の針列  $\xi = \{v_n\}$  に対して指示列  $\{I(x, \sigma, \xi)_n\}$  は指示関数  $I(x)$  に収束する；
- (A.2) 任意に固定した正数  $\epsilon$  に対して

$$\sup_{\text{dist}(x, \Sigma) > \epsilon} I(x) < \infty$$

が成り立つ；

- (A.3)  $\text{Int } \Sigma$  上の任意に与えられた点  $a$  に対して

$$\lim_{x \rightarrow a} I(x) = \infty$$

が成り立つ。

この定理は探針法の A 面について述べたものであるが当然次の疑問が浮かぶ。針の先端が亀裂上に到達した場合あるいは針が亀裂を貫通した場合、指示列はどのように振舞うか。

これに答えるのが次の定理である。

**定理 6.1.B.**  $x$  は  $\Omega \setminus \partial\Sigma$  内の任意の点、 $\sigma$  は  $x$  を先端にもつ針で、 $\sigma([0, 1])$  は  $\Sigma$  と共有点をもちかつそのすべてが  $\text{Int } \Sigma$  に含まれると仮定する。そのとき  $(x, \sigma)$  に対する任意の針列  $\xi = \{v_n\}$  に対して  $\lim_{n \rightarrow \infty} I(x, \sigma, \xi)_n = \infty$  が成り立つ。

### 6.3. 反射解の爆発

与えられた  $\Omega$  全体での調和関数  $v \in H^1(\Omega)$  に対して、 $u \in X(\Omega \setminus \Sigma; D)$  は  $f$  が  $v$  の  $\partial\Omega$  上のトレースで与えられるときの (6.1) の弱解とする。そのとき関数  $w = u - v \in X_0(\Omega \setminus \Sigma; D)$  は  $D$  の選び方によらない。この  $w$  を  $v$  の  $\Sigma$  による反射解と呼ぶ。

定理 6.1.A および B の証明は次のなんの変哲もない  $w$  の特徴づけから出発する。証明は演習問題とする。

**補題 6.2.**  $v$  の  $\Sigma$  による反射解  $w$  は次の等式を満たす：

$$\int_{\Sigma} \frac{\partial v}{\partial \nu} (\Psi_+ - \Psi_-) dS = \int_{\Omega \setminus \Sigma} \nabla w \cdot \nabla \Psi dy, \forall \Psi \in X_0(\Omega \setminus \Sigma; D). \quad (6.5)$$

次の命題がこの節の核心部分であり、 $v$  の  $\sigma$  による反射解のエネルギーの、 $v$  のみを使った下からの評価を与えている。証明は弱解の枠内で初等的になされる。



命題 6.3.  $\eta \in C_0^\infty(\Omega)$  および正の定数  $M$  は

$$\|\eta\|_{L^\infty(\Omega_-)} + \|\nabla\eta\|_{L^\infty(\Omega_-)} \leq M \quad (6.6)$$

および

$$\text{supp}(\eta|_\Gamma) \subset \Sigma \quad (6.7)$$

を満たすとせよ。 $\Omega_-$  と  $\text{supp} \eta$  との共通部分を  $(\Omega_-)_\eta$  と書く。 $v$  は  $H^1(\Omega)$  に属する調和関数および  $w \in X_0(\Omega \setminus \Sigma; D)$  は  $v$  の  $\Sigma$  による反射解とする。もし

$$\int_{\Omega_-} |\nabla v|^2 dy - \left| \int_\Gamma \frac{\partial v}{\partial \nu} (1 - \eta) \bar{v} dS \right| \geq 0 \quad (6.8)$$

ならば、そのとき反射解のエネルギーについて次の評価が成り立つ：

$$\frac{(\int_{\Omega_-} |\nabla v|^2 dy - \left| \int_\Gamma \frac{\partial v}{\partial \nu} (1 - \eta) \bar{v} dS \right|)^2}{M^2 (\int_{(\Omega_-)_\eta} |\nabla v|^2 dy + \int_{(\Omega_-)_\eta} |v|^2 dy)} \leq \|\nabla w\|_{L^2((\Omega_-)_\eta)}^2. \quad (6.9)$$

証明要は (6.5) のうまい試験関数  $\Psi$  を構成することであるが、それは次のように、 $v$  と  $\eta$  から作られる：

$$\Psi(y) = \begin{cases} 0, & y \in \Omega_+, \\ -\eta(y)\bar{v}(y), & y \in \Omega_-. \end{cases}$$

$\Psi$  の  $\partial\Omega$  へのトレースは消えていることおよび  $\Gamma$  上、等式  $\Psi_+ - \Psi_- = \eta\bar{v}$  が成り立つことは明らかである。条件 (6.7) は、この  $\Psi$  が  $X_0(\Omega \setminus \Sigma; D)$  の要素であることを保証している。(6.6) からその  $\Omega_-$  におけるエネルギーの  $v$  による評価

$$\|\nabla \Psi\|_{L^2(\Omega_-)}^2 = \|\nabla(\eta\bar{v})\|_{L^2(\Omega_-)}^2 \leq M^2 (\int_{(\Omega_-)_\eta} |\nabla v|^2 dy + \int_{(\Omega_-)_\eta} |v|^2 dy) \quad (6.10)$$

を得る。さて部分積分によれば等式

$$\begin{aligned} \int_{\Omega_-} |\nabla v|^2 dy &= \int_\Gamma \frac{\partial v}{\partial \nu} \bar{v} dS = \int_\Gamma \frac{\partial v}{\partial \nu} \eta \bar{v} dS + \int_\Gamma \frac{\partial v}{\partial \nu} (1 - \eta) \bar{v} dS \\ &= \int_\Sigma \frac{\partial v}{\partial \nu} (\Psi_+ - \Psi_-) dS + \int_\Gamma \frac{\partial v}{\partial \nu} (1 - \eta) \bar{v} dS \end{aligned}$$

を得る。この右辺第一項に (6.5) を適用して評価

$$\int_{\Omega_-} |\nabla v|^2 dy \leq \|\nabla w\|_{L^2((\Omega_-)_\eta)} \|\nabla \Psi\|_{L^2((\Omega_-)_\eta)} + \left| \int_\Gamma \frac{\partial v}{\partial \nu} (1 - \eta) \bar{v} dS \right| \quad (6.11)$$

を得る。(6.10) と (6.11) より目的の評価 (6.9) を得る。

□

針列  $\xi = \{v_n\}$  の各項  $v_n$  とその勾配  $\nabla v_n$  との関係は明らかではない。次の補題は  $\nabla v_n$  が大きくなる場所では  $v_n$  は  $\nabla v_n$  と比べてそれほど大きくないということを表現している。

補題 6.4.  $x$  は  $\Omega$  内の任意に与えられた点、 $\sigma$  は  $x$  を先端にもつ任意の針、 $\xi = \{v_n\}$  は  $(x, \sigma)$  に対する任意の針列とする。もし

$$\lim_{n \rightarrow \infty} \int_{\Omega_-} |\nabla v_n|^2 dy = \infty$$

ならば、そのとき十分大きい自然数  $n_0$  を取ると、数列

$$\left\{ \frac{\int_{\Omega_-} |v_n|^2 dy}{\int_{\Omega_-} |\nabla v_n|^2 dy} \right\}_{n \geq n_0}$$

は有界である。

証明記述の簡単のため、 $\Omega_- = D$  が一つの連結成分からなる場合のみ考えよう。 $\mathbf{R}^2$  のコンパクト集合で  $\Omega \setminus \sigma([0, 1])$  に含まれる集合からなる列  $\{K_l\}$  で条件  $K_l \subset K_{l+1}$  for  $l = 1, \dots$  および  $\Omega \setminus \sigma([0, 1]) = \bigcup_{l=1}^{\infty} K_l$  を満たすものをとる。そのとき、Lebesgue の単調収束定理および  $\sigma([0, 1])$  は測度 0 であるという事実から、 $l \rightarrow \infty$  のとき  $|K_l \cap \Omega_-|$  は  $|\Omega_- \setminus \sigma([0, 1])| = |\Omega_-|$  に収束する。したがって十分大きい  $l_0$  をとると、集合  $A \equiv K_{l_0} \cap \Omega_-$  は正の測度をもつ。Poincaré の不等式によれば評価

$$\begin{aligned} \int_{\Omega_-} |v_n|^2 dy &\leq 2 \int_{\Omega_-} |v_n - (v_n)_A|^2 dy + 2 \int_{\Omega_-} |(v_n)_A|^2 dy \\ &\leq 2C(\Omega_-, A)^2 \int_{\Omega_-} |\nabla v_n|^2 dy + 2|\Omega_-| |(v_n)_A|^2 \end{aligned}$$

を得る。ここで  $C(\Omega_-, A)$  は  $v_n$  に独立な正の定数、 $(v_n)_A$  は  $v_n$  の  $|A|$  における平均値である：

$$(v_n)_A = \frac{1}{|A|} \int_A v_n dy.$$

$\bar{A}$  は  $\Omega \setminus \sigma([0, 1])$  に含まれるから、数列  $\{(v_n)_A\}$  は収束する。これと上の不等式から証明は終わる。

□

定理 6.1.B の証明を述べよう。要点は最初に選んだ  $D$  をその境界  $\Gamma$  が補題 6.2 の (6.8) を満たすようにとりなおすことである。その指針は次のとおりである。針が、 $\text{Int } \Sigma$  と初めて当たる点以前に  $\Gamma \setminus \text{Int } \Sigma$  と共有点が存在しないとき。このときは  $\Gamma \setminus \text{Int } \Sigma$  の部分を大きく変形してそれと  $\Sigma$  によって囲まれた領域  $D'$  でその境界  $\Gamma'$  と針の共通部分が  $\text{Int } \Sigma$  のみに含まれるようにする。針が、 $\text{Int } \Sigma$  と初めて当たる点以前に  $\Gamma \setminus \text{Int } \Sigma$  と共有点をもっているときは、 $\Gamma \setminus \text{Int } \Sigma$  を避けて  $\Sigma$  に沿って  $\Omega_+$  の側に、 $\Gamma \setminus \text{Int } \Sigma$  を折り返して押し込んでそれと  $\Sigma$  によって囲まれた領域  $D'$  をとれば最初の場合に帰着される。

このとり直しと反射解の  $D$  に関する不変性から、はじめから、 $\sigma([0, 1])$  と  $\Gamma$  との共通部分は  $\text{Int } \Sigma$  に含まれるとしてよい。集合  $\sigma([0, 1]) \cap \Gamma$  はコンパクトであるから、関数  $\eta \in C_0^\infty(\Omega)$  で  $\sigma([0, 1]) \cap \Gamma$  のある近傍で恒等的に 1 に等しくかつ (6.7) を満たすものがとれる。命題 3.4 の (3.3) を  $v = \partial v_n / \partial y_j - \partial G(\cdot - x) / \partial y_j$  および  $\Omega \setminus \sigma([0, 1])$  にその閉包がふくまれる任意の開円板  $W$  に適用して、高階微分の収束

$$v_n \rightarrow G(\cdot - x) \text{ in } H_{\text{loc}}^2(\Omega \setminus \sigma([0, 1]))$$

を得る。集合  $\{y \in \Gamma \mid \eta(y) \neq 1\}$  は  $\Omega \setminus \sigma([0, 1])$  に含まれるから、上で述べた収束性とトレース定理により数列

$$\left\{ \int_{\Gamma} \frac{\partial v_n}{\partial \nu} (1 - \eta) \bar{v}_n dS \right\}$$

は有界であることが分かる。一方補題 3.3 および 3.5 によれば発散

$$\lim_{n \rightarrow \infty} \int_{\Omega_-} |\nabla v_n|^2 dy = \infty \quad (6.12)$$

を得る。したがって

$$\lim_{n \rightarrow \infty} \frac{\int_{\Gamma} \frac{\partial v_n}{\partial \nu} (1 - \eta) \bar{v}_n dS}{\int_{\Omega_-} |\nabla v_n|^2 dy} = 0 \quad (6.13)$$

を結論する。したがって特に十分大きい  $n_0$  に対して  $n \geq n_0$  を満たすすべての  $n$  に対して (6.8) が満たされることになる。適当に正数  $M$  をとれば (6.6) が満たされることは明らかであり、補題 6.4 によれば、適当な正数  $K$  をとるとすべての  $n$  に対して

$$\int_{\Omega_-} |\nabla v_n|^2 dy + \int_{\Omega_-} |v_n|^2 dy \leq K \int_{\Omega_-} |\nabla v_n|^2 dy \quad (6.14)$$

が成り立つことになる。(6.12), (6.13), (6.14) および (6.9) から  $w_n$  の  $(\Omega_-)_\eta$  におけるエネルギーの爆発をみることは容易である。後は、(6.4) を使って指示列の爆発を得る。

□

定理 6.1.A の (A.3) の証明も同様にして出来るので省略。(A.2) および (A.3) は順問題の適切性に関係しており演習問題とする (ルーチンという意味で)。

$\Sigma$  が十分滑らかと仮定すると  $\Sigma$  による反射解  $w_x(y)$  を  $\text{Int } \Sigma$  上の与えられた点の十分小さい近傍で、 $x \rightarrow a$  のときそのエネルギーの発散に寄与する項を  $y$  の関数としてとることができる。これを使って (A.3) を証明することができる。それは鏡像の方法の一般化と考えられる。ただしその方法は、例えば弾性体の方程式系のように方程式が連立になると、格段に複雑になる。また B 面については無力のように見える。これらの詳細については [46, 64] を参照されたい。とにかく、ここで展開した方法は弾性体の方程式系へも全く同じように適用できるであろう。なお因数分解法 [57] のアイデアを用いた論文 [7] が存在することを注意しておく。

三次元物体の中に発生した、未知の一枚の平面上に乗っている未知の亀裂という特殊な幾何を仮定した場合においては、**適当な条件を満たす一組の電位と電流密度分布**を用いた Andrieux-Abda[1] の再構成公式がある。

## 7 囲い込み法と逆散乱問題

### 7.1. 体積のない障害物による音波の逆散乱問題

前節までは、すべて Laplace 方程式にかかわる逆問題であった。音波、電磁波あるいは弾性波にかかわる逆問題は Helmholtz 方程式  $\Delta u + k^2 u = 0$  に帰着される問題が多くある。ここでは、体積のない、音響的に固い障害物による音波の逆散乱問題の二次元版をとりあげ、単純化された囲い込み法の最新の応用を述べよう。

ここで考える逆問題は無限遠方から平面波を入射しその障害物による反射波の無限遠方での漸近形から障害物の位置および形についての情報を抽出することである。

ただし以下の三つの条件を考慮する：

- 入射平面波の波数は固定して、限りなく大きくも（高周波）、限りなく小さくも（低周波）しないということ；

- 入射平面波の入射方向は有限個であること；

- 漸近形を見る方向もできれば限定すること（視野限定）。

これらは応用を考える上で避けてはならない条件であり、これをどう克服するかそのアイデアが興味のあるところである。

さて散乱問題を述べるのであるが、障害物（の断面）を  $\Sigma$  であらわすことにし、それは有限個の互いに素な折れ線  $\Sigma_1, \Sigma_2, \dots, \Sigma_m$  で与えられていると仮定する。適当に関数空間を設定し、変分法的定式化と Fredholm の交代定理を使うと（例えば [22] を参照せよ）、与えられた正数  $k$  および入射方向  $d \in S^1$  に対して次の散乱問題の解が一意的に存在することがわかる：

- $u$  は Helmholtz 方程式

$$\Delta u + k^2 u = 0 \text{ in } \mathbf{R}^2 \setminus \Sigma$$

を満たす；

- 

$$\frac{\partial u}{\partial \nu} = 0 \text{ on } \Sigma;$$

- $w = u - e^{ikx \cdot d}$  は Sommerfeld の放射条件

$$\lim_{r \rightarrow \infty} \sqrt{r} \left( \frac{\partial w}{\partial r} - ikw \right) = 0$$

を満たす。ここに  $r = |x|$ 。

● 折れ線の折れ曲がっているところや端点での  $u$  の正則性は高々  $H^1$  になっている。  
 $w$  は障害物による反射波をあらわしている。このとき与えられた方向  $\varphi \in S^1$  に対して  $r \rightarrow \infty$  のとき  $w$  は漸近展開

$$w(r\varphi) = \frac{e^{ikr}}{\sqrt{r}} F(\varphi; d, k) + O\left(\frac{1}{r^{3/2}}\right)$$

を持つことを示すことができる。この展開にあらわれる  $F(\varphi; d, k)$  は  $w(r\varphi)$  の **far field pattern** と呼ばれる。これがこの節の観測データである。

これを使って問題を述べよう。

**逆問題 7.1.**  $\Gamma$  は任意に与えられた  $S^1$  の空でない部分集合とする。与えられた有限個の  $d$  および  $k$  に対して、観測データ  $\{F(\varphi; d, k) \mid \varphi \in \Gamma\}$  から  $\Sigma$  の位置および形についての情報を抽出せよ。

この問題はまさしく典型的逆問題であろう。

## 7.2. 散乱振幅から凸包を抽出すること

関数

$$h_\Sigma(\omega) = \sup_{x \in \Sigma} x \cdot \omega, \quad \omega \in S^1$$

を  $\Sigma$  の支持関数という。方向  $\omega \in S^1$  の  $\Sigma$  に関する正則性は定義 2.3 の  $D$  を  $\Sigma$  に置き換えて定義される。この節では方向  $\varphi = (\varphi_1, \varphi_2) \in S^1$  と対応する複素数  $\varphi_1 + i\varphi_2$  を同じ記号  $\varphi$  であらわす。この規約を使って、与えられた  $k > 0$ 、 $N = 1, \dots$ 、 $\tau > 0$  および  $\omega \in S^1$  に対して  $S^1$  上で定義された関数

$$g_N(\varphi; \tau, k, \omega) = \frac{1}{2\pi} \sum_{|m| \leq N} \left\{ \frac{ik\varphi}{(\tau + \sqrt{\tau^2 + k^2})\omega} \right\}^m \quad (7.1)$$

を考える。次の定理は、逆問題 7.1 の  $\Gamma = S^1$  の場合 (全視野) における一つの解答を与える。

**定理 7.1.**  $\Sigma$  は、原点を中心とするある既知の半径  $R$  の開円板  $B_R$  にふくまれていると仮定する。方向  $\omega$  は  $\Sigma$  に関して正則とせよ。 $\beta_0$  は次の方程式の正の解とする：

$$2s + e \log s = 0.$$

$\beta_0$  より小さい正数  $\beta$  を任意に固定する。正数からなる数列  $\{\tau(N)\}_{N=1, \dots}$  で  $N \rightarrow \infty$  のとき

$$\tau(N) = \frac{\beta N}{eR} + O(1) \quad (7.2)$$

を満たすものを任意に与える。二つの一次独立な方向  $d_1$  および  $d_2$  を任意に与える。そのとき、 $N \rightarrow \infty$  のとき

$$\frac{\log\left\{\sum_{j=1}^2 \left| \int_{S^1} F(-\varphi; d_j, k) g_N(\varphi; \tau(N), k, \omega) dS(\varphi) \right| \right\}}{\tau(N)} \rightarrow h_\Sigma(\omega) \quad (7.3)$$

が成り立つ。さらに以下の主張が成り立つ：

- もし  $t \geq h_\Sigma(\omega)$  ならば、そのとき

$$\lim_{N \rightarrow \infty} e^{-\tau(N)t} \sum_{j=1}^2 \left| \int_{S^1} F(-\varphi; d_j, k) g_N(\varphi; \tau(N), k, \omega) dS(\varphi) \right| = 0;$$

- もし  $t < h_\Sigma(\omega)$  ならば、そのとき

$$\lim_{N \rightarrow \infty} e^{-\tau(N)t} \sum_{j=1}^2 \left| \int_{S^1} F(-\varphi; d_j, k) g_N(\varphi; \tau(N), k, \omega) dS(\varphi) \right| = \infty.$$

公式 (6.3) の意味するところは、もし反射波の Far field pattern の Fourier 係数

$$\int_{S^1} F(\varphi; d_j, k) \varphi^m dS(\varphi) = (-1)^m \int_{S^1} F(-\varphi; d_j, k) \varphi^m dS(\varphi), \quad |m| \leq N, j = 1, 2$$

が十分大きい  $N$  に対して分かっているならば、そのとき次の計算式

$$\frac{\log\left\{\sum_{j=1}^2 \left| \int_{S^1} F(-\varphi; d_j, k) g_N(\varphi; \tau(N), k, \omega) dS(\varphi) \right| \right\}}{\tau(N)}$$

が  $h_\Sigma(\omega)$  の近似になっていることを保証するというものである。(7.2) をよくみると次のような示唆を得るであろう： $R$  が大きいときすなわち  $\Sigma$  の評価が粗雑なときは、 $\tau(N)$  を大きくするには、 $N$  をずっと先までとることが必要であるだろう； $R$  が小さいときすなわち  $\Sigma$  の評価が良いときは  $N$  がそんなに大きくなくても  $\tau(N)$  は大きくなり、(7.3) の左辺はよく  $h_\Sigma(\omega)$  を近似するであろう。

しかし手離しでは喜べないこともわかる。それは前提である (7.2) という  $\tau(N)$  についての条件であるが、これはあくまでも漸近式であり、それを満たす  $\tau(N)$  は無数にあるとい

うことである。これは結局は  $N$  が有限のとき、(7.1) の  $\tau$  をどうとるかという問題であり、理論としてこれ以上のことを言うのは難しい問題であろう。

### 7.3. Herglotz の波動関数、Vekua 変換および指示関数の漸近展開

定理 7.1 の証明自体が面白いので概略を述べよう。最初の補題は反射波の Far field pattern の表現公式である (例えば [12])。

**補題 7.2.**  $S^1$  上の通常の測度に関して可測な任意の部分集合  $\Gamma$  を与える。 $\mathbb{R}^2 \setminus B_R$  で滑らかな関数  $u$  は  $\mathbb{R}^2 \setminus \overline{B}_R$  で Helmholtz 方程式  $\Delta u + k^2 u = 0$  を満たしかつ  $w = u - e^{ikx \cdot d}$  は Sommerfeld の放射条件

$$\lim_{r \rightarrow \infty} \sqrt{r} \left( \frac{\partial w}{\partial r} - ikw \right) = 0$$

を満たすとする。そのとき公式

$$\int_{\Gamma} F(-\varphi; d, k) g(\varphi) dS(\varphi) = -\frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\partial B_R} \left( \frac{\partial u}{\partial \nu} v_g - \frac{\partial v_g}{\partial \nu} u \right) dS$$

が成り立つ。ここで  $v_g$  は密度  $g \in L^2(\Gamma)$  をもつ Herglotz 波動関数：

$$v_g(y) = \int_{\Gamma} e^{iky \cdot \varphi} g(\varphi) dS(\varphi)$$

および  $\nu$  は  $B_R$  に対する外向き単位法線ベクトル場である。

次の補題は (7.1) が何故出てくるのかを説明している。

**補題 7.3.** 任意に与えられた方向  $\omega = (\omega_1, \omega_2)$  に対して  $\omega^\perp = (\omega_2, -\omega_1)$  とおく。定理 7.1 の中の関数  $g = g_N(\cdot; \tau(N), k, \omega)$  は、 $y \in \overline{B}_R$  に関して一様に (導関数まで込めて)、 $N \rightarrow \infty$  のとき漸近式

$$\int_{S^1} e^{iky \cdot \varphi} g(\varphi) dS(\varphi) = e^{y \cdot (\tau(N)\omega + i\sqrt{\tau(N)^2 + k^2}\omega^\perp)} + O(e^{-R\tau(N)} N^{-\infty})$$

を満たす。

この補題の証明は、次の Vekua 変換 [74] を使うと組織的に理解される：

$$T_k v(x) = v(x) - \frac{k|x|}{2} \int_0^1 v(tx) J_1(k|x|\sqrt{1-t}) \frac{dt}{\sqrt{1-t}}$$

ここで  $v$  は任意の調和関数、 $J_1$  は一次の Bessel 関数である。各  $m = 0, 1, 2, \dots$  に対して  $T_k$  は性質

$$T_k : r^m e^{\pm im\theta} \mapsto \left(\frac{2}{k}\right)^m m! J_m(kr) e^{\pm im\theta}$$

をもつことは、容易にわかる。ここで  $J_m(z)$  は  $m$  位の Bessel 関数である：

$$J_m(z) = \left(\frac{z}{2}\right)^m \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n+1+m)} \left(\frac{z}{2}\right)^{2n}.$$

この性質に注意すれば  $g$  を見つける問題は近似式

$$\int_{S^1} \{e^{ik\bar{\varphi}(y_1 + iy_2)/2} + e^{ik\varphi(y_1 - iy_2)/2} - 1\} g(\varphi) dS(\varphi) \approx e_\omega(y; \tau, k)$$

を満たす  $g$  を見つける問題に帰着される。ここで  $e_\omega(y; \tau, k)$  は調和関数で  $T_k\{e_\omega(\cdot; \tau, k)\} = e^{y \cdot (\tau\omega + i\sqrt{\tau^2 + k^2}\omega^\perp)}$  を満たすもので、具体的に書けるが省略する。

そして結局、 $g_N(\varphi; \tau, k, \omega)$  は積分方程式

$$\int_{S^1} \{e^{ik\varphi(y_1 + iy_2)/2} + e^{ik\varphi(y_1 - iy_2)/2} - 1\} g(\varphi) dS(\varphi) = e_\omega(y; \tau, k)$$

の形式解である発散級数

$$g(\varphi) = \sum_{m=0}^{\infty} \frac{1}{2\pi} \left\{ \frac{ik\varphi}{(\tau + \sqrt{\tau^2 + k^2})\omega} \right\}^m + \sum_{m=1}^{\infty} \frac{1}{2\pi} \left\{ \frac{(\tau + \sqrt{\tau^2 + k^2})\omega}{ik\varphi} \right\}^m$$

を打ち切って得られる。 $N$  に対する  $\tau(N)$  の選び方 (6.2) は Stirling の公式に由来するが詳細は省略する。

さて最も中心となる補題を述べよう。これは単純化された囲い込み法における定理 2.6 と補題 2.7 に対応する。その証明は劇的に難しそうであったが (1999 年当時)、後述の波数に依存したパラメタの変更というアイデアで切り抜けることができた (2003 年夏の ICIAM2003 から帰国して数日)。

**補題 7.4.**  $d = d_j$  に対する  $u$  を  $u_j$  と書き、

$$v = e^{x \cdot (\tau\omega + i\sqrt{\tau^2 + k^2}\omega^\perp)}, \tau > 0$$

とおく。方向  $\omega$  は  $\Sigma$  に関して正則であるとせよ。そのとき  $\tau$  に独立な適当な正数  $\mu$  および  $A$  をとると、 $\tau \rightarrow \infty$  のとき

$$\tau^\mu e^{-\tau h_\Sigma(\omega)} \sum_{j=1}^2 \left| \int_{\partial B_R} \left( \frac{\partial u_j}{\partial \nu} v - \frac{\partial v}{\partial \nu} u_j \right) dS \right| \rightarrow A$$

が成り立つ。

証明の概要を述べよう。

最初に次の積分の  $\tau \rightarrow \infty$  における漸近挙動を調べる：

$$I_j(\tau) = \int_{\partial B_R} \left( \frac{\partial u_j}{\partial \nu} v - \frac{\partial v}{\partial \nu} u_j \right) dS.$$

部分積分により等式

$$e^{-\tau h_\Sigma(\omega)} I_j(\tau) = e^{-\tau h_\Sigma(\omega)} \int_\Sigma [u_j] \frac{\partial v}{\partial \nu} dS \quad (7.4)$$

を得る。ここで  $[u_j] = u_j^+|_\Sigma - u_j^-|_\Sigma$  は  $u$  の  $\Sigma$  の両側からのトレースの差である (詳しくは述べないが、 $\nu$  の向いている方を  $-$  にとった)。

集合  $\{x \mid x \cdot \omega = h_\Sigma(\omega)\} \cap \Sigma$  は一点からなるので、その点を  $x_0$  であらわそう。このとき命題 2.5 に対応する  $u$  の  $x_0$  における展開公式が成り立つことがわかる。そしてパラメタの変更

$$s = \sqrt{\tau^2 + k^2} + \tau \quad (7.5)$$

と (7.4) を使うと、 $s \rightarrow \infty$  のとき完全漸近展開

$$e^{-i\sqrt{\tau^2 + k^2}x_0 \cdot \omega^\perp} e^{-\tau h_\Sigma(\omega)} I_j(\tau) \sim \sum_{l=1}^{\infty} \frac{A_{l,j}}{s^{\lambda_l}} \quad (7.6)$$

を得る。ここで  $0 < \lambda_1 < \lambda_2 < \cdots \rightarrow \infty$  および  $\{A_{l,j}\}_l$  は  $u_j$  の  $x_0$  における展開係数および  $\Sigma$  の  $x_0$  の近傍における幾何を含んでいる。この部分の計算は込み入っているが、しかし計算する喜びをもたらしてくれた。この (7.5) の変更が本質的であるのは (7.6) の各項をみればあきらかであろう。

次に (7.6) の展開係数のなかに消えないものが本当にあるのか否かを調べるのであるが、直接はできず、背理法により以下の場合にわけられる：

- もし  $x_0$  がどの  $\Sigma_1, \dots, \Sigma_m$  の端点になっていなければ、すなわち、 $\Sigma$  がそこで折れ曲がっている点であるならば、そのとき各  $j = 1, 2$  に対して  $A_{l,j} \neq 0$  となる  $l = l_j$  が存在する；

- もし  $x_0$  がある  $\Sigma_p$  の端点であるならば、そのときすべての  $j = 1, 2$  およびすべての  $l$  に対して  $A_{l,j} = 0$  となることはない。この証明において入射方向  $d_1, d_2$  が一次独立であることが本質的にきいている。

これらをまとめると補題 7.4 が成り立つことがわかる。

定理 7.1 はこれら補題を使えば直ちに証明される。実際補題 7.2 より等式

$$\begin{aligned} & -\tau(N)^\mu e^{-\tau(N)h_\Sigma(\omega)} \frac{\sqrt{8\pi k}}{e^{i\pi/4}} \times \int_{S^1} F(-\varphi; d_j, k) g_N(\varphi; \tau(N), k, \omega) dS(\varphi) \\ & = \tau(N)^\mu e^{-\tau(N)h_\Sigma(\omega)} \int_{\partial B_R} \left( \frac{\partial u_j}{\partial \nu} v - \frac{\partial v}{\partial \nu} u_j \right) dS \\ & + \tau(N)^\mu e^{-\tau(N)h_\Sigma(\omega)} \int_{\partial B_R} \left\{ \frac{\partial u_j}{\partial \nu} (v_{g_N} - v) - \frac{\partial}{\partial \nu} (v_{g_N} - v) u_j \right\} dS \end{aligned}$$

を得る。ここで

$$v = e^{y \cdot (\tau(N)\omega + i\sqrt{\tau(N)^2 + k^2}\omega^\perp)}$$

および  $v_{g_N}$  は密度  $g = g_N(\cdot; \tau(N), k, \omega)$  に対する Herglotz 波動関数である。  $h_\Sigma(\omega) \geq -R$  に注意すれば補題 7.3 より右辺第二項は  $N \rightarrow \infty$  のとき減衰する。あとは補題 7.4 より、 $N \rightarrow \infty$  のとき

$$\tau(N)^\mu e^{-\tau(N)h_\Sigma(\omega)} \sqrt{8\pi k} \sum_{j=1}^2 \left| \int_{S^1} F(-\varphi; d_j, k) g_N(\varphi; \tau(N), k, \omega) dS(\varphi) \right| \rightarrow A$$

を得る。定理はこれからただちに導かれる。

#### 7.4. 視野が限定されたときの先験情報の役割

(7.3) の中には反射波の far field pattern のすべての方向における値があらわれている。では見る方向が一部に限定されたいとき（向こう側から視えないとき）、そのような公式があるかという疑問が当然浮かぶ。これは、限定された視野をもつ far field pattern が観測データとして与えられた場合である。ここではある先験情報を取り入れることにより、定理 7.1 が局所化できることを示そう。以下では、 $S^1$  の空でない開集合  $\Gamma$  を任意に与え固定する。ここでも  $\Sigma$  は、原点を中心とするある既知の半径  $R$  の開円板  $B_R$  にふくまれていると仮定する。

Helmholtz 方程式の解からなる集合

$$W(B_R) = \{v \in C^2(B_R) \cap C^1(\overline{B_R}) \mid \Delta v + k^2 v = 0 \text{ in } B_R\}$$

を導入する。  $W(B_R)$  の  $H^1(B_R)$  における閉包を  $\overline{W(B_R)}$  によってあらわそう。任意に与えられた  $g \in L^2(-\Gamma)$  に対して

$$Hg(y) = \int_{-\Gamma} e^{iky \cdot \varphi} g(\varphi) dS(\varphi), \quad y \in B_R.$$



により  $B_R$  上の関数を定義する。そのとき  $Hg \in \overline{W(B_R)}$  でありかつ作用素  $H : L^2(-\Gamma) \rightarrow \overline{W(B_R)}$  は有界である。次の事実が知られている ([13])。

**定理 7.5.**  $H$  の像は  $\overline{W(B_R)}$  で稠密である。

さて与えられた  $v \in \overline{W(B_R)}$  および正数  $\delta$  に対して、 $L^2(-\Gamma)$  の要素  $g_0$  が積分方程式  $Hg = v$  の食い違いが  $\delta$  の最小ノルム解であるとは、 $g_0$  が不等式  $\|Hg - v\|_{H^1(B_R)} \leq \delta$  を満たしかつ等式

$$\|g_0\|_{L^2(-\Gamma)} = \inf\{\|g\|_{L^2(-\Gamma)} : \|Hg - v\|_{H^1(B_R)} \leq \delta\}$$

が成り立つときを言う。

特に与えられた正数  $\tau$  および方向  $\omega \in S^1$  に対して  $v = e^{x \cdot (\tau\omega + i\sqrt{\tau^2 + k^2}\omega^\perp)}$  ととる。これはもちろん  $\overline{W(B_R)}$  の要素である。このとき、一般論によれば ([59])、定理 7.5 は任意に固定した正数  $\delta$  に対して、積分方程式  $Hg = v$  の食い違いが  $\delta$  の最小ノルム解の存在を保証する。しかもそれは公式

$$g = (\alpha I + H^*H)^{-1}H^*v$$

で与えられる。ここで  $\alpha$  は

$$\|H(\alpha I + H^*H)^{-1}H^*v - v\|_{H^1(B_R)}^2 = \delta^2.$$

を満たす正の数である。したがって  $\alpha$  は  $\tau, \omega, k, \delta$  に依存する。そこでこの最小ノルム解を  $g = g_{\tau, \delta}(\cdot; k, \omega)$  とあらわそう。これはもちろん不等式

$$\|Hg_{\tau, \delta}(\cdot; k, \omega) - v\|_{H^1(B_R)} \leq \delta \quad (7.7)$$

を満たしている。そのとき次の定理を得る。

**定理 7.6.**  $\omega$  は  $\Sigma$  に関して正則で、 $h_\Sigma(\omega) > 0$  を仮定する。そのとき公式

$$\lim_{\tau \rightarrow \infty} \frac{\log \sum_{j=1}^2 \left| \int_{\Gamma} F(\varphi; d_j, k) g_{\tau, \delta}(\varphi; k, \omega) dS(\varphi) \right|}{\tau} = h_\Sigma(\omega),$$

が成り立つ。さらに以下の主張は正しい:

- もし  $t \geq h_\Sigma(\omega)$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} e^{-\tau t} \sum_{j=1}^2 \left| \int_{\Gamma} F(\varphi; d_j, k) g_{\tau, \delta}(\varphi; k, \omega) dS(\varphi) \right| = 0;$$

- もし  $t < h_D(\omega)$  ならば、そのとき

$$\lim_{\tau \rightarrow \infty} e^{-\tau t} \sum_{j=1}^2 \left| \int_{\Gamma} F(\varphi; d_j, k) g_{\tau, \delta}(\varphi; k, \omega) dS(\varphi) \right| = \infty.$$

証明トレース定理を使えば、(7.7) から評価

$$\|Hg_{\tau, \delta} - v\|_{H^{1/2}(\partial B_R)} + \left\| \frac{\partial}{\partial \nu} \{Hg_{\tau, \delta} - v\} \right\|_{H^{-1/2}(\partial B_R)} \leq C\delta \quad (7.8)$$

を得る。ここで  $C$  は  $\tau$  に独立な正数である。補題 7.2 によれば、等式

$$\begin{aligned}
& -e^{-\tau h_{\Sigma}(\omega)} \frac{\sqrt{8\pi k}}{e^{i\pi/4}} \int_{\Gamma} F(\varphi; d_j, k) g_{\tau, \delta}(\varphi; k, \omega) dS(\varphi) \\
& = e^{-\tau h_{\Sigma}(\omega)} \int_{\partial B_R} \left( \frac{\partial u_j}{\partial \nu} v - \frac{\partial v}{\partial \nu} u_j \right) dS + e^{-\tau h_{\Sigma}(\omega)} \int_{\partial B_R} \left\{ \frac{\partial u_j}{\partial \nu} (H g_{\tau, \delta} - v) - \frac{\partial}{\partial \nu} (H g_{\tau, \delta} - v) u_j \right\} dS
\end{aligned} \tag{7.9}$$

を得る。仮定  $h_{\Sigma}(\omega) > 0$  および (7.8) は  $\tau \rightarrow \infty$  のとき (7.9) の右辺第二項の指数的減衰を導く。あとは前と同じように補題 7.4 から定理は従う。

□

$h_{\Sigma}(\omega) > 0$  は座標の原点と  $\Sigma$  との位置関係に関する条件である。観測データの不足をこの条件で救っているのがみてとれるであろう。この定理の数値実験は大変興味がある。 $g_{\tau, \delta}$  は精度よく計算できるのであろうか。

この節で展開した方法は、音響的に硬い体積のある障害物や、背景とは異なる密度を持った、体積のある障害物にたいする類似の問題の二次元版にもすでに適用できることがわかっている。[39, 42] を参照されたい。しかし、三次元の問題についてはどうか、あるいは**弾性波の亀裂による散乱問題**においてどうかというのは、最も興味がある今後の問題である。

## 謝辞

日本学術振興会科学研究費補助金 (C)(2) (No. 15540154) の援助に感謝する。広島大学の池島良氏との長年にわたるさまざまな議論は研究を進める上で励みになったことを記しておきたい。そして共同研究者で囲い込み法の数値実験を粘り強く進めてくれた、GE medical system の Samuli Siltanen 氏および岡山理科大学の大江貴司氏に深く感謝する。日本応用数理学会および北大 COE にはこのような機会を与えて頂いたことに感謝する。特に東京大学の奈良高明氏には、原稿を準備するにあたっての、著者の提起した問題点を解消するため動いていただき、存分に執筆に専念することができた。深く感謝する。

## References

- [1] Andrieux, S. and Abda, B. A., Identification of planar cracks by complete over-determined data: inversion formulae, *Inverse Problems*, 12(1996), 553-563.
- [2] Alessandrini, G. and DiBenedetto, E., Determining 2-Dimensional cracks in 3-dimensional bodies: uniqueness and stability, *Indiana Univ. Math. J.*, 46(1997), 1-82.
- [3] Bateman, H., Higher Transcendental Functions, Volume III, Bateman Manuscript Project (A. Erdélyi, Editor), New York, McGRAW-HILL, 1955.
- [4] Bal, G., Reconstructions in impedance and optical tomography with singular interfaces, *Inverse Problems*, 21(2005), 113-131.

- [5] Brühl, M., Explicit characterization of inclusions in electrical impedance tomography, *SIAM J. Math. Appl.*, 32(2001), 1327-1341.
- [6] Brühl, M., Hanke, M., Numerical implementation of two noniterative methods for locating inclusions by impedance tomography, *Inverse Problems*, 16 (2000), 1029-1042.
- [7] Brühl, M., Hanke, M. and Pidcock, M., Crack detection using electrostatic measurements, *Mathematical Modelling and Numerical Analysis*, 35(2001), 595-605.
- [8] Calderón, A. P., On an inverse boundary value problem, in *Seminar on Numerical Analysis and its Applications to Continuum Physics* (Meyer, W. H. and Raupp, M. A. eds.), Brazilian Math. Society, Rio de Janeiro, 1980, 65-73.
- [9] Cheng, J., Liu, J. J. and Nakamura, G., Recovery of the shape of an obstacle and the boundary impedance from the far-field pattern, *J. Math. Kyoto Univ.*, 43(2003), 165-186.
- [10] Cheng, J., Liu, J. J. and Nakamura, G., The numerical realization of the probe method for the inverse scattering problems from the near field data, *Inverse Problems*, 21(2005), 839-855.
- [11] Colton, D. and Kirsch, A., A simple method for solving inverse scattering problems in the resonance region, *Inverse Problems*, 12(1996), 383-393.
- [12] Colton, D. and Kress, R., Eigenvalues of the far field operator for the Helmholtz equation in an absorbing medium, *SIAM J. Appl. Math.*, 55(1995), 1724-1735.
- [13] Colton, D. and Kress, R., On the denseness of Herglotz wave functions and electromagnetic Herglotz pairs in Sobolev spaces, *Math. Meth. Appl. Sci.*, 24(2001), 1289-1303.
- [14] Grinberg, N., Obstacle visualization via the factorization method for the mixed boundary value problem, *Inverse Problems*, 18(2002), 1687-1704.
- [15] Grinberg, N. and Kirsch, K., The linear sampling method in inverse obstacle scattering for impedance boundary conditions, *J. Inv. Ill-Posed Problems*, 10(2002), 171-185.
- [16] Daido, Y., Ikehata, M. and Nakamura, G., Reconstruction of inclusions for the inverse boundary value problem with mixed type boundary condition, *Appl. Anal.*, 83(2004), 109-124.
- [17] Eller, M., Identification of cracks in three-dimensional bodies by boundary measurements, *Inverse Problems*, 12(1996), 395-408.
- [18] Erhard, K. and Potthast, R., A numerical study of the probe method, submitted.
- [19] Friedman, A. and Isakov, M., On the uniqueness in the inverse conductivity problem with one measurements, *Indiana Univ. Math. J.*, 38(1989), 563-579.

- [20] Friedman, A. and Vogelius, M., Determining cracks by boundary measurements, *Indiana Univ. Math. J.*, 38(1989), 527-556.
- [21] Grisvard, P., *Elliptic problems in nonsmooth domains*, Pitman, Boston, 1985.
- [22] Hettlich, F., On the uniqueness of the inverse conductive scattering problem for the Helmholtz equation, *Inverse Problems*, 10(1994), 129-144.
- [23] Hyvönen, N., Characterizing inclusions in optical tomography, *Inverse Problems*, 20(2004), 737-752.
- [24] Ikehata, M., Reconstruction of the shape of the inclusion by boundary measurements, *Comm. PDE.*, 23(1998), 1459-1474.
- [25] Ikehata, M., Reconstruction of an obstacle from the scattering amplitude at a fixed frequency, *Inverse Problems*, 14(1998), 949-954.
- [26] Ikehata, M., Reconstruction of obstacle from boundary measurements, *Wave Motion*, 30(1999), 205-223.
- [27] Ikehata, M., Enclosing a polygonal cavity in a two-dimensional bounded domain from Cauchy data, *Inverse Problems*, 15(1999), 1231-1241.
- [28] Ikehata, M., Reconstruction of a source domain from the Cauchy data, *Inverse Problems*, 15(1999), 637-645.
- [29] Ikehata, M., Reconstruction of the support function for inclusion from boundary measurements, *J. Inv. Ill-Posed Problems*, 8(2000), 367-378.
- [30] 池畠 優, 不連続面の観測データによる再構成の数値-境界値逆問題, 逆ソース問題および逆散乱問題における再構成公式, 都立大学数学教室セミナー報告, 2000.
- [31] Ikehata, M., On reconstruction in the inverse conductivity problem with one measurement, *Inverse Problems*, 16(2000), 785-793.
- [32] Ikehata, M., On reconstruction from a partial knowledge of the Neumann-to-Dirichlet operator, *Inverse Problems*, 17(2001), 45-51.
- [33] Ikehata, M., Inverse conductivity problem in the infinite slab, *Inverse Problems*, 17(2001), 437-454.
- [34] Ikehata, M., Exponentially growing solutions and the Cauchy problem, *Appl. Anal.*, no. 1-2, 78(2001), 79-95.
- [35] Ikehata, M., Reconstruction of inclusion from boundary measurements, *J. Inv. Ill-Posed Problems*, 10(2002), 37-65.
- [36] Ikehata, M., Extracting the convex hull of an unknown inclusion in the multilayered material, *Appl. Anal.*, 82(2003), 857-873.
- [37] Ikehata, M., Complex geometrical optics solutions and inverse crack problems, *Inverse Problems*, 19(2003), 1385-1405.

- [38] Ikehata, M., Exponentially growing solutions, multilayered anisotropic material and the enclosure method, *Inverse Problems*, 19(2003), 1065-1079.
- [39] Ikehata, M., Inverse scattering problems and the enclosure method, *Inverse Problems*, 20(2004), 533-551.
- [40] Ikehata, M., Mittag Leffler's function and extracting from Cauchy data, *Inverse problems and spectral theory*, Isozaki, H. Ed., *Contemporary Math.*, 348(2004), 41-52.
- [41] Ikehata, M., The Herglotz wave function, the Vekua transform and the enclosure method, submitted, September, 2004.
- [42] Ikehata, M., An inverse transmission scattering problem and the enclosure method, *Computing*, 2005.
- [43] Ikehata, M., A new formulation of the probe method and related problems, *Inverse Problems*, 21(2005), 413-426.
- [44] Ikehata, M., Two sides of probe method and obstacle with impedance boundary condition, submitted, March, 2005.
- [45] Ikehata, M., Inverse crack problem and probe method, to appear, *Cubo*.
- [46] Ikehata, M. and Nakamura, G., Reconstruction formula for identifying cracks, *J. Elasticity*, 70(2003), 59-72.
- [47] Ikehata, M. and Ohe, T., Numerical method for finding the convex hull of polygonal cavities using enclosure method, *Inverse Problems*, 18(2002), 111-124.
- [48] Ikehata, M. and Ohe, T., A numerical method for finding the convex hull of inclusions using the enclosure method, in *Electromagnetic nondestructive evaluation (VI)*, Kojima, F. et al. (Eds.), IOS Press, Amsterdam, 2002, 21-28.
- [49] Ikehata, M. and Siltanen, S., Numerical method for finding the convex hull of an inclusion in conductivity from boundary measurements, *Inverse Problems*, 16(2000), 1043-1052.
- [50] Ikehata, M. and Siltanen, S., Electrical impedance tomography and Mittag-Leffler's function, *Inverse Problems*, 20(2004), 1325-1348.
- [51] Ikehata, M. and Siltanen, S., Numerical solution of the Cauchy problem for the stationary Schrödinger equation using Faddeev's Green function, *SIAM J. Appl. Math.*, 64(2004), 1907-1932.
- [52] Isakov, V., On uniqueness of recovery of a discontinuous conductivity coefficients, *Comm. Pure. Appl. Math.*, 41(1988), 865-877.
- [53] Kang, H., Lim, M. and Nakamura, G., Detection of surface breaking cracks in two dimensions, *Inverse Problems*, 19 (2003), 909-918.

- [54] Kirsch, A., Characterization of the shape of a scattering obstacle using the spectral data of the far field operator, *Inverse Problems*, 14(1998), 1489-1512.
- [55] Kirsch, A., Factorization of the far-field operator for the inhomogeneous medium case and an application in inverse scattering theory, *Inverse Problems*, 15(1999), 413-429.
- [56] Kirsch, A., New characterizations of solutions in inverse scattering theory, *Appl. Anal.*, 76(2000), 319-350.
- [57] Kirsch, A. and Ritter, S., A linear sampling method for inverse scattering from an open arc., *Inverse Problems*, 16(2000), 89-105.
- [58] Kozlov, V. A., Mazy'ya, V. G. and Rossmann, J., Spectral problems associated with corner singularities of solutions to elliptic equations, *Mathematical Surveys and Monographs*, 85(2000), AMS.
- [59] Kress, R., *Linear Integral equations*, Springer, 1989.
- [60] Lele, A. S., Kulkarni, S. R. and Willsky, A. S., Convex-polygon estimation from support-line measurements and applications to target reconstruction from laser-radar data, *J. Opt. Soc. Am. A*, 9(1992), 1693-1714.
- [61] Ladyzhenskaya, O. A. and Ural'tzeva N. N., *Linear and quasilinear elliptic equations*, 1968, London, Academic Press.
- [62] Moriya, K. and Endo, T., A study of flaw detection method for CFRP composite laminates (1st Report). The measurement of crack extension in CFRP composites by electrical potential method (in Japanese), *Journal of the Japan Society for Aeronautical and Space Science*, 36(1988), 139-146.
- [63] Nakamura, G. and Wang, J. N., Unique continuation for an elasticity system with residual stress and its applications, *SIAM J. Math. Anal.*, 35 (2003), 304-317.
- [64] Nakamura, G., Uhlmann, G. and Wang, J. N., Reconstruction of cracks in an inhomogeneous anisotropic elastic medium, *J. Math. Pures Appl.*, 82(2003), 1251-1276.
- [65] Nakamura, G., Uhlmann, G. and Wang, J. N., Oscillating-decaying solutions, Runge approximation property for the anisotropic elasticity system and their applications to inverse problems, *J. Math. Pures Appl.*, 84 (2005), 21-54.
- [66] Potthast, R., Stability estimates and reconstructions in inverse scattering using singular sources, *J. Comp. Appl. Math.*, 114(2000), 247-274.
- [67] Potthast, R., Sampling and probe methods-an algorithmical view, *Computing*, 2005.
- [68] Prince, J. J. and Willsky, A. S., Reconstructing convex sets from support line measurements, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(1990), 377-389.
- [69] Sylvester, J. and Uhlmann, G., A global uniqueness theorem for an inverse boundary value problem, *Ann. Math.*, 125(1987), 153-169.

- [70] 田辺広城、関数解析 上、実教出版、1978.
- [71] Tilo, A., Why linear sampling works, *Inverse Problems*, 20(2004), 163-173.
- [72] Todoroki, A., Kobayashi, H and Matsuura, K., Application of electrical potential method as delamination sensor for smart CFRP structures (in Japanese), *Journal of the Japan Society for Composite Materials*, 21(1995), 89-95.
- [73] Uhlmann, G., Developments in inverse problems since Calderón's foundational paper, in *Harmonic analysis and partial differential equations* (Christ, M., Kenig, C. E. and Sadosky, C., eds.), 1999, pp. 295-345, The University of Chicago Press, Chicago and London.
- [74] Vekua, I., Modification of an integral transformation and some of its properties, *Bull. Acad. Sci. Georgian SSR*, 6(1945), 177-183.
- [75] Yarmukhamedov, Sh., On the Cauchy problem for the Laplace equation, in Russian, *Dokl. Akad. Nauk SSSR*, 235(1977), 281-283.
- [76] Yosida, K., *Functional Analysis*, Third Edition, Springer, 1971.

e-mail address

ikchata@math.sci.gunma-u.ac.jp

## 数値計算の視点から：多倍長計算の逆問題、非適切問題への応用

藤原 宏志 (京都大学大学院 情報学研究科)

fujiiwara@acs.i.kyoto-u.ac.jp

逆問題の数学解析と同じく、数値解析理論と数値計算技法も重要である。本稿では、逆問題の数値計算手法から、「多倍長計算」「スペクトル法」の基礎を紹介する。

順問題を対象とする数値解析と数値計算は、系の安定性を前提としている。系が数値的に安定とは、誤差の混入があったとしても、その誤差が微小であれば計算結果への影響は少なく、数値計算が信頼できることを意味する。安定性を有する問題では、分割数を増やすことで数値解の精度の向上が期待され、順問題に対する大規模数値計算は、計算機の性能の向上とともに成功を収めてきた。これはいわば、数値計算の「量の改善」である。

さて逆問題とひとことでいっても、領域決定問題、方程式の未知係数、ソース項や初期値の決定問題など様々な問題設定がなされる。数値解法では個々の系がもつ特徴を保存するような離散化が望ましく、従って問題に応じた議論が必要となる。しかしこれら逆問題の数値計算にはある共通する数学的性質があり、数値計算を困難なものにしている。この困難性は逆問題の非適切性 (ill-posedness) に起因するものであり、**数値的不安定性 (numerical instability)** と呼ばれる。

数値計算では、実数とその演算、微分作用素や積分作用素の近似をおこなうため、誤差が混入する。従って、数学的に正しい式が与えられたとしても、それを計算機上で厳密 (exact) に実現することは殆んど望めない。数値的に不安定な問題の数値計算では、これらの誤差が急激に増大して、計算が破綻する。そのため「量の改善」だけでは逆問題の数値計算は実現し得ないことが指摘されている。

数値的不安定性の克服のために、最小二乗法、一般逆法、正則化法などの理論がある。これらは微分方程式などで記述される系から最小化問題を導出し、その最小化問題を解くものである。数値計算にあたっては、最小化問題を離散化して有限個の未知数 (パラメータ) 探索問題に帰着させる。これらの方法は汎用性が高いという利点をもつ反面、その裏返しとして、導出された最小化問題では元の問題や解のもつ特徴的な意味や重要な性質が失われていることも多い。未知パラメータの個数や値に人工的な制限を加えることもあり、元の問題の特性を欠落した議論も少なくなかった。

最小化問題に話を限ると、近年は勾配法のみならず、AI、ニューラルネットや遺伝的アルゴリズムにより最適パラメータ探索をおこなう数値計算も試みられている。計算機の高性能化や逆問題に対する知識の蓄積もあり、これらの手法が効果的である場合もある。また何らかの要請で数値計算が必要なとき、これらの手法に頼らざるを得ないというのも事実であるが、これらの手法の数学的な正当化は困難である。

この反省のもと、近年、逆問題の数値的不安定性を克服する新たな手法が幾つか提案されている。そのひとつが、**計算機上での実数の扱いの改善と高精度離散化の併用という、数値計算の「質の改善」**である。この手法により、数値的不安定性の克服の可能性が示されつつある。これらの手法を具体的・現実的な逆問題に適用していくことが、逆問題解析の数値解析と数値計算における課題である。未完成の手法でもあり、またこれらの手法だけでは不十分でもある。今後は、具体的な適用の中での数値解析と計算手法の進展が望まれる。



## 1 計算機上での実数の近似と丸め誤差

電子計算機で扱うことが可能な情報 (データ) は, 本質的には 2 値の区別, すなわち “0” と “1” のみである. この最小単位をビット (binary unit) とよぶ. ただし 2 値をとるビットの扱いのみでは不便なので, 実際には「複数のビットを束ねて」それらに「型 (type)」の概念を付与することにより, 整数, 実数, 文字, 画像や音などの多様なデータを扱うことが可能となる.

計算機上での数値計算において実数を表現する型は複数提案されている. 科学技術計算で広範に利用されているのは浮動小数点数とよばれる型であり, 演算などのアルゴリズムとあわせて浮動小数点方式として定義される. 浮動小数点方式の一般論については, [7, 16] に詳しい.

多倍長数, あるいは多倍長計算の明確な定義はない. プロセサのアーキテクチャおよびプログラミング言語で定義される基底型に比して, 多くの精度を有する実数近似手法とその演算と考えるのが自然である. 本稿でもこれに従い, 多倍長精度・有限桁の浮動小数点方式について述べる. そこで基礎となるのは浮動小数点方式の理論である. 本節では浮動小数点数のデータ構造について述べる.

### 浮動小数点数 $F_3$

浮動小数点方式は, 実数の表現方法のひとつである. 例えば 10 進法 5 桁の浮動小数点方式では 123.45 を  $+1.2345 \times 10^2$  のように表す.  $+$  は符号であり, 10 を基数 (base または radix), 1.2345 を仮数部 (significand), 10 の冪である 2 を指数とよぶ. 浮動小数点方式では表現が一意になるように, 仮数部の先頭桁を 0 以上基数未満のひと桁の整数で表すのが一般的である.

浮動小数点数は基数, 仮数部の桁数 (精度), 指数の範囲によっていくつかの方法が提案されている. 現在ではワークステーションやパーソナルコンピュータでは 10 進法 (基数が 10) は使われず, 2 進法 (基数 2) が利用される.

以下の浮動小数点数の集合  $F_3$  を例に, 典型的な 2 進法の浮動小数点数のデータ構造を説明する.

$$\begin{aligned} \pm f_0.f_1f_2 \times 2^e, \\ f_i \in \{0, 1\}, \quad e \in \{\perp, -1, 0, 1, \top\}. \end{aligned} \tag{1.1}$$

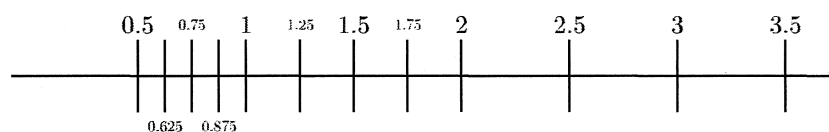
$e$  が指数,  $f_0.f_1f_2$  が仮数部である.  $F_3$  の元は, 指数部が 5 つの状態をとり, 仮数部の精度が 3 桁 (3 ビット) で構成される.

まず,  $e$  が  $-1, 0, 1$  のいずれかである場合には,  $f_0 = 1$  であると約束する. したがって (1.1) が表す数のうち, 正のものを全て列挙すると

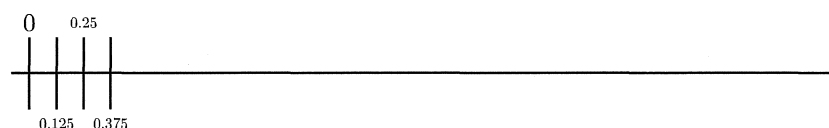
$1.00 \times 2^{-1} = 0.5$	$1.00 \times 2^0 = 1$	$1.00 \times 2^1 = 2$
$1.01 \times 2^{-1} = 0.625$	$1.01 \times 2^0 = 1.25$	$1.01 \times 2^1 = 2.5$
$1.10 \times 2^{-1} = 0.75$	$1.10 \times 2^0 = 1.5$	$1.10 \times 2^1 = 3$
$1.11 \times 2^{-1} = 0.875$	$1.11 \times 2^0 = 1.75$	$1.11 \times 2^1 = 3.5$

となる. ここで, 左辺の仮数部は 2 進法表現であり, 右辺は 10 進法表現である. これらを実数直線上に表したものが図 1(a) である.  $e \in \{-1, 0, 1\}$  で表されるこれらの数を正規化数 (normalized number) と呼ぶ.  $F_3$  では, 0 は正規化数では表現できないことに注意する.

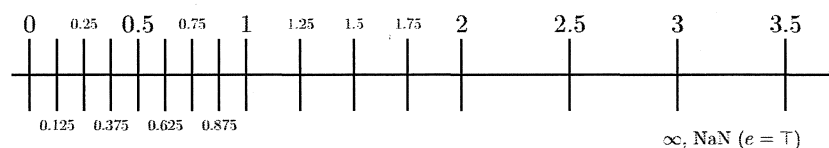
次に  $e = \perp$  の場合を考える. このときは  $f_0 = 0$  と約束し, (1.1) の  $e$  には正規化数を表現する指数部の最小値を代入するものと約束する.  $F_3$  では,  $2^\perp = 2^{-1}$  である.  $e = \perp$  で表現される数



(a)  $F_3$  の正規化数



(b)  $F_3$  の非正規化数



(c)  $F_3$  で表現される正数

図 1: 浮動小数点数の体系  $F_3$

のうち正のものを列挙すると,

$$0.00 \times 2^{-1} = 0$$

$$0.01 \times 2^{-1} = 0.125$$

$$0.10 \times 2^{-1} = 0.25$$

$$0.11 \times 2^{-1} = 0.375$$

となり, 数直線上では図 1(b) のように並ぶ.  $e = \perp$  の場合に表されるこれらの数を非正規化数 (denormalized number) とよぶ. 近年は subnormalized number とよばれることも多い.

$e = \top$  の場合は  $2^\top$  の値を考えず, 演算結果として  $\pm\infty$  であることを表したり, 演算とその結果が無効であることを示す not a number (NaN) の状態を示す.

このとおり,  $F_3$  で表される数には正規化数と非正規化数がある. 正の数を合わせて数直線上に表すと, 図 1(c) のようになり, 浮動小数点数は実数軸上に離散的に存在することがわかる. また, 等間隔には並ばないことも特徴である. となりあう 2 数の差は仮数部  $f_0.f_1f_2$  の最下位桁における誤差に相当する. この差のことを ulp (units in the last place) とよぶ.

## 丸め

浮動小数点数は実数軸上に離散的に存在するため、すべての実数を厳密に表すことはできない。となりあう 2 つの浮動小数点数の間にある実数は、含まれる区間の両端のうち一方の浮動小数点数で近似して計算機上で表現する。この近似のことを丸めという。いいかえると、丸めは、実数から浮動小数点数の集合への写像である。ある実数を丸める際には、その実数が含まれる (最小の) 区間の両端のうち、実数に近いほう (等距離のときはいずれか一方) を採る最近値丸め (rounding to the nearest), 大きいほうを採る  $+\infty$  方向への丸め (rounding up), 小さいほうを採る  $-\infty$  方向への丸め (rounding down), 0 に近いものを選択する 0 方向への丸め (chopping) の 4 種類が考えられる。通常は最近値丸めがおこなわれる。

最近値丸めでは、扱いたい実数とそれを丸めて得られる浮動小数点数との間に最大で  $\frac{1}{2}\text{ulp}$  の (絶対) 誤差が含まれる。この誤差を丸め誤差 (rounding error) という。また、 $\frac{1}{2}\text{ulp}$  に相当する相対誤差を計算機イプシロンということが多く、この場合、実数の表現における丸め誤差は計算機イプシロン以下になる [7]。この計算機イプシロンの定義は決まったものではなく、例えば [13] では  $1\text{ulp}$  に相当する相対誤差を計算機イプシロンと定義している。

## 浮動小数点数 $F_3$ のコーディング

$F_3$  を計算機で扱う際に必要なビット数を考えよう。符号は正か負のいずれかであるから、1 ビットで表現可能である。指数部の 5 つの状態を表現するには 3 ビットあれば充分である。仮数部については、 $f_0$  は  $e$  の状態から決定されるため省略し、 $f_1f_2$  のみの表現に 2 ビットでよい。このとき省略された  $f_0$  は、暗黙の省略の 1 ビット (hidden bit) あるいは、暗黙の省略の 1 (hidden one) とよばれる。合わせると、 $F_3$  の浮動小数点数 (1.1) の保持には 6 ビットあればよいことがわかる。実装の際には符号、指数、仮数部をそれぞれどのようなビット列に割当てても決めなければならないが、ここでは省略する。

絶対値が最小・最大の数のオーダーは指数部のとり得る値によって決定され、となりあう 2 数の距離は仮数部の桁数により決定される。従って、より広い範囲の実数を扱いたい場合には指数部を増やせばよい。また、丸め誤差を小さくする、すなわち隣り合う 2 つの浮動小数点数の間隔を小さくしたい場合には仮数部の精度を増やせばよい。

## IEEE754 倍精度数

一般的な数値計算における実数の扱いは IEEE754 で定義される方式に準拠している。IEEE754 では単精度方式、倍精度方式などが定義されており、それぞれ浮動小数点数の表現に際しては表 1 に示すビット数が割当てられている。特にその中でも、倍精度方式 (double format) が数値計算で標準的に利用されている。倍精度方式は 64 ビットの情報によって図 2 のように計算機の内部では保持される。倍精度数は符号部  $s$ 、バイアスされた指数部  $e_b (= e + 1023)$ 、仮数部  $F$  から構成される。 $e_b \neq 0, 1023$  の場合が正規化数にあたり、図 2 によって

$$(-1)^s \times 2^{e_b-1023} \times (1 + F \times 2^{-52}) \quad (1.2)$$

という数を表すこととする。ここで、図 2 の  $e_b, F$  は 2 進法で符号無し整数と解釈する。 $e_b = 0, 1023$  の場合も含めて、倍精度数は表 2 のようなコーディングがなされる。ただし  $0x$  で始まる数字は 16 進法による記述である。丸めには、先に説明した 4 つのモードがある。最近値丸めにおいて最近値が 2 つ存在する場合には、結果の最下位ビットが 0 である方を選ぶ偶数丸めが採用されている。

表 1: Floating Point Format Parameters Defined in IEEE754

Parameter	Format			
	Single		Double	Double Extended
	Single	Extended		
Format width in bits	32	$\geq 43$	64	$\geq 79$
Exponent width in bits	8	$\geq 11$	11	$\geq 15$
Significand width in bits (precision)	24	$\geq 32$	53	$\geq 64$

表 2: IEEE754 倍精度数で表現される数と指数部のコーディング

指数部のコーディング ( $e_b$ )	指数部 ( $e$ )	仮数部 ( $F$ )	浮動小数点数値 ( $x$ )	magnitude
0x000	$-0x3fe$ ( $\perp$ )		$\pm 0.F \times 2^{-1022}$	subnormal
0x001	$-0x3fe = -1022$		$\pm 1.F \times 2^{-1022}$	$10^{-300}$
0x002	$-0x3fd = -1021$		$\pm 1.F \times 2^{-1021}$	
$\vdots$	$\vdots$		$\vdots$	
0x3fe	$-1$		$\pm 1.F \times 2^{-1}$	
0x3ff	$0$		$\pm 1.F$	$1 \dots$
0x400	$1$		$\pm 1.F \times 2$	$2 \dots$
$\vdots$	$\vdots$		$\vdots$	
0x7fd	$0x3fe = 1022$		$\pm 1.F \times 2^{1022}$	$10^{300}$
0x7fe	$0x3ff = 1023$		$\pm 1.F \times 2^{1023}$	
0x7ff	$\top$	$0$	$\pm \infty$	$\pm \infty$
0x7ff	$\top$	$\neq 0$	Not a Number (NaN)	

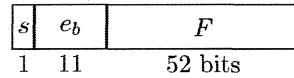


図 2: IEEE754 倍精度数の計算機の内部での表現

## 丸め誤差

浮動小数点方式では、どのような方法を採用にせよ、遅延評価などの手法を利用せずに精度を固定する限りは、実数表現には丸め演算が不可欠である。この近似は実数の表現の際におこるだけではない。浮動小数点数の集合は通常の実数の四則演算について閉じていない。さらに浮動小数点数値に対する初等関数の値も、一般には浮動小数点数に含まれない。そのため、各種の演算結果を浮動小数点数値で得るたびに、丸め演算が必要となる。

表現と演算結果の丸めに際して混入するこれらの誤差を丸め誤差と呼ぶのであるが、特に演算においては「桁落ち現象」、「情報落ち現象」と呼ばれる計算機特有の現象を生じる。

**桁落ち現象 (truncation)** 「有効数字の上位桁が等しい 2 数の加減算の結果、有効数字が減少する現象」。

例えば、10 進法で 5 桁の精度を有する浮動小数点数体系  $DF_5$  において演算を考えた場合、

$$1.2953 \times 10^2 - 1.2951 \times 10^2 = 2.0000 \times 10^{-2}$$

となる。このとき被減数  $1.2953 \times 10^2$  および減数  $1.2951 \times 10^2$  は 5 桁の有効数字を保持するが、演算結果である 2 数の差  $2.0000 \times 10^{-2}$  の有効数字は 1 桁のみである。ここで、演算結果の小数点以下に続く 4 つの 0 は桁をあわせるためだけに詰められるだけのものであり、有効数字ではないことに注意する。桁落ちでは誤差は混入しないが、桁落ちに続く演算により大きな誤差の混入を招く可能性がある。

**情報落ち現象 (loss of trailing digits)** 「極端に絶対値が異なる 2 数の加算もしくは減算により、一方の情報が計算結果に反映されない現象」。

上述の 10 進 5 桁の  $DF_5$  では、次の演算がおこなわれる。

$$1.2953 \times 10^2 + 9.2331 \times 10^{-10} = 1.2953 \times 10^2$$

この式における等号  $=$  は、 $DF_5$  での演算結果を意味する。この例では  $9.2331 \times 10^{-10}$  を加えているにも関わらず、予め定められた桁数の制限によりこの数の情報は計算結果に反映されていない。

2つの現象で注意すべきことは、**浮動小数点有限桁方式では数学的に厳密な式が与えられたとしても、計算機上で正確な結果を得られるとは限らない**という事実である。計算過程に現れる全ての(浮動小数点方式で近似された)実数や数式に対し、桁落ちと情報落ちが生じないように式変型をおこなうことは困難であり、丸め誤差の処理に加えて、これらの計算機現象に常に留意する必要がある。

## 計算環境に依存する丸め誤差

現在、一般的に普及しているプロセサは浮動小数点方式として IEEE754 に準拠していることが多いが、使用する OS、コンパイラ、ライブラリなどによって数値計算結果が異なることがある。基本的な演算での相違を表 3.4 に示す。また、組み込み関数の値も異なる場合がある(表 5.[17])。

表 3: 計算環境による計算結果 1 (倍精度)

$x = 2^{-53}, y = 2^{-54}$	$1 + x + y$	$1 + (x + y)$
Xeon (Linux) Pentium(Linux, Borland C++ Cygwin, MinGW)	$1 + 2^{-52}$	$1 + 2^{-52}$
Alpha, UltraSPARC-II Itanium2, PowerPC Pentium (FreeBSD, Visual C++)	1	$1 + 2^{-52}$

表 4: 計算環境による計算結果 2 (倍精度)

$$a \approx 0.854200011227730815477 \dots$$

$$b \approx 0.519944555523513818329 \dots$$

$a^2 - b^2$	0.459315318362911 1073 ...
Alpha, SPARC, Itanium2 Pentium(FreeBSD)	0.459315318362911 06 ...
Pentium(Linux)	0.459315318362911 12 ...

表 5: 計算環境による組み込み関数の計算値 (倍精度)

$x$	64	$2^{46} (\approx 7.04 \times 10^{13})$
$\sin(x)$	0.9200260381967906 83...	0.9994730524837994 98...
SPARC	0.9200260381967907 4...	0.9994730524837994 7...
Alpha		
PowerPC G4	0.9200260381967906 3...	0.9994730524837994 7...
P4(Cygwin)		
Xeon(Linux)		
P4 (Linux, FreeBSD)	0.9200260381967906 3...	0.9994730495431699 6...
BC++, VC++, MinGW)		

これらの表中の数値は計算機上で 2 進法で得られた結果を 10 進法で記述しており、17 桁以降を省略して ... としているために正確な演算結果を示しているものではないことを注意する。2 進法の浮動小数点演算を厳密に表記するためには C99 規格 [15] で導入された記法を利用すると便利である。詳しくは述べないが、 $\pm 0 \times 1.F p e$  という表記により実数  $\pm(1 + F \times 2^{-52}) \times 2^e$  を表す。 $F(0 \leq F < 2^{52})$  は 16 進法整数で、 $e$  は 10 進法で記述する。これらは C 言語プログラムの中でも有効な定数句である。

この例のように、個々の表現や演算に伴う丸め誤差は小さいため、これだけを見ていると大きな問題ではないように思える。しかし前述のとおり、計算機アーキテクチャによる浮動小数点方式の僅かな違いが計算結果が大きく影響を与える場合がある。

例えば、次の三項間漸化式を考えよう。

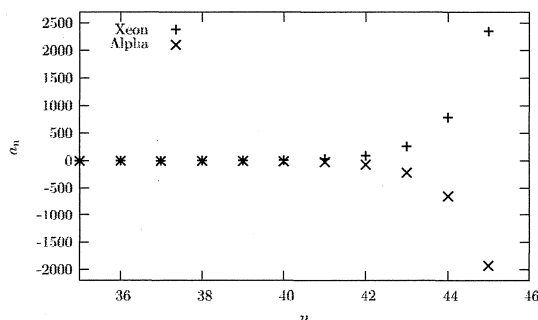
$$a_{n+2} = \frac{34}{11}a_{n+1} - \frac{3}{11}a_n,$$

$$a_0 = 1, \quad a_1 = \frac{1}{11}.$$

この漸化式の  $\{a_n\}$  の値を C 言語による数値計算で求めると、図 3 のように、利用する計算機によって著しく異なる数値計算結果が得られる。さらに、いずれの数値計算結果も厳密解  $a_n = (1/11)^n$  とは全く異なっていることに注意する。実はこの例の漸化式の一般解は  $a_n = c_1(1/11)^n + c_2 3^n$  であり、 $3^n$  の発散項を有するためにその数値計算は数値的に不安定である。このため、この漸化式を与えられた式のとおりに計算すると、丸め誤差のために  $c_2 \neq 0, |c_2| \ll 1$  の場合の計算に相当する可能性がある。個々の演算での丸め誤差は僅かであるが、数値的に不安定な問題では累積された丸め誤差が著しく増大し、図 3 のように利用する計算環境に依存して計算結果に深刻な影響を与える。

## 2 多倍長計算環境 exflib

多倍長数の標準的な表現と実現方式は存在せず、著者は独自に多倍長精度・浮動小数点数の型 exfloat (extended precision float-int-point number) を定義し、その演算を exflib (extended precision floating-point arithmetic library) として実装している。exfloat 型の浮動小数点数は仮数部の桁数は可変長で、ユーザが自由に決定することができる。すなわち隣り合う 2 つの浮動小数点数の距離を任意に変更することが可能であり、丸め誤差の大きさを任意に設定することができる。標語で



$n$	Xeon(Linux)	Alpha(Linux)
2	0.00826446	0.00826446
5	$6.20921 \times 10^{-6}$	$6.20921 \times 10^{-6}$
10	$3.86014 \times 10^{-11}$	$3.85159 \times 10^{-11}$
40	9.68369	-7.91327
50	571812	-467270

図 3: Numerically Unstable Processes

いうと, exflib は数値計算の世界における顕微鏡であり, これまでの数値計算では見えなかったものを計算機上で実現することを目指としている. 本節では, exfloat, exflib とその演算について概説する.

exflib は, 科学技術計算, 特に偏微分方程式や積分方程式などの数値計算での利用を想定して, 10 進法で 100 桁から数千桁程度の精度での演算を, 特に演算の高速性, メモリの効率的な利用を念頭において設計している. 現段階の exflib は C++ 言語からの利用が可能で, 浮動小数点数の四則演算, 組み込み関数値, 10 進法による初期化と出力など, 数値計算に必要な機能を有している.

exflib は C 言語や FORTRAN から利用可能な演算サブルーチンの集まりであり, ライブラリとして提供している. このライブラリを使いやすくするため, C++ 言語におけるインターフェースも提供している. 多倍長精度での計算が実行可能な環境としては多くの実装があり, 現在は GNU MP, Pari, MPFR, bc が代表的である. また国内では UBASIC が著名である. 商用の数式処理ソフトである Maple, Mathematica でも利用可能である.

### 多倍長数 exfloat のデータ構造

exflib では多倍長精度の浮動小数点数 exfloat を定義し, メモリ中に図 4 のように 64 ビット符号なし整数の配列により保持しており,

$$(-1)^s \times 2^{e_b-b} \times \left(1 + \sum_{i=1}^n f_i \times 2^{-64i}\right)$$

という実数を表している. ただし,  $b = 2^{62} - 1$  である. 64 ビット計算機での利用を考えてこの実装を採っているが, 32 ビット計算機においてもメモリに格納する最には 64 ビット配列として格納している. そのため, MPI などを利用して 64 ビット計算機と 32 ビット計算機が混在するような環境においても並列計算が可能である. 仮数部を倍精度数で実装する環境もあるが, 64 ビット計算機では精度と計算時間の両面において整数型を基底型として利用するほうがよい. これにより, メモリの効率的な利用およびメモリアクセスの回数の減少という利点もあわせもつ.

仮数部は, 配列に記憶された  $64n$  ビットに加えて暗黙の省略の 1 ビットをもつ. 従って exfloat 型の浮動小数点数は  $64n + 1$  ビットの精度をもつ. この  $n$  はユーザが自由に指定することができるが, 現在のところ, ひとつのプログラム中では精度の変更などはできない. 演算アルゴリズムの都合上  $n \geq 3$  (10 進法では 77 桁以上に相当) を仮定している. 四則演算のアルゴリズムでは  $n$  に上限は設定していないが, 組み込み関数で利用する数表は  $n = 1024$  (約 2 万桁) の精度を有する.



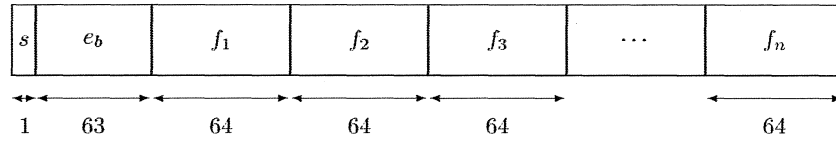


図 4: 多倍長精度・浮動小数点数 exfloat 型のフォーマット

指数部には固定長の 63 ビットを割当てており、 $10^{-10^{18}}$  から  $10^{10^{18}}$  までの数を扱うことが可能である。ただしプログラムの簡素化と演算の高速化のために非正規化数は定義しておらず、 $e_b = 0$  のときは仮数部の値によらず常に 0 と解釈する。倍精度方式で非正規化数として扱われる範囲の数も exfloat では正規化数として扱われるため、一般的な数値計算で扱われる範囲の実数は exfloat で表現することが可能であると考えている。 $e_b = 2^{63} - 1$  の場合は  $\pm\infty$  として扱う。

#### 浮動小数点演算のアルゴリズム

exflib は exfloat に対する演算を提供している。浮動小数点数の演算アルゴリズムについては [8]、多倍長演算については [16] が詳しい。理論的な誤差解析については [7] に詳しい。

#### exflib の加減算のアルゴリズム

加減算では古典的な筆算による方法を実装している。加減算における保護桁は 3 ビットで充分であるが [7]、プログラムの簡素化のために exflib ではアーキテクチャのビット数に応じて切りがよいように、余分に保護桁をとっている。

$n$  ビットアーキテクチャでは  $n$  桁の符号なし整数の加算において結果は  $n + 1$  桁となる。すなわち、繰上りを考慮にいれる必要がある。C 言語や FORTRAN ではこの繰上りを捕捉する命令はないが、大抵のアーキテクチャでは特別なレジスタが定義されているため、これを利用すれば繰上りの発生を捕捉することができ、特別なアルゴリズムを必要としない。Alpha アーキテクチャなどではこのような命令が定義されておらず、同等のアルゴリズムを実装する必要がある [6]。C 言語や FORTRAN で実現する場合にはそのようなアルゴリズムを利用するほかに、インライン・アセンブラを利用したり、精度の低い基底型を利用する方法が考えられる。

#### exflib の乗算の演算アルゴリズム

乗算の古典的アルゴリズムは筆算である。筆算における基底型の乗算の回数は精度の 2 乗に比例する。一般にプロセッサ内部での整数型の乗算は加減算に比して 10 倍程度の時間がかかるとされており [4, 14]、基底型の乗算処理がアルゴリズム全体の計算時間の大部分を占める。乗算は組み込み関数などでも多用される重要な演算であり、基底型の乗算の回数を減らすためのアルゴリズムの研究が盛んにおこなわれている [16]。

代表的なものは Karatsuba-Offman のアルゴリズムである。基数を  $\beta$  とし、2 つの数  $a\beta + b, c\beta + d$

の筆算を考える。ただし  $0 \leq a, b, c, d < \beta$  である。この2数の積は次のようになる。

$$(a\beta + b)(c\beta + d) = ac\beta^2 + (ad + bc)\beta + bd \quad (2.1)$$

$$= ac\beta^2 + ((a-b)(c-d) - ac - bd)\beta + bd. \quad (2.2)$$

式 (2.2) が Karatsuba のアルゴリズムを示しており、式中の  $ac, (a-b)(c-d), bd$  の積もそれぞれ Karatsuba 法により再帰的に計算する。

式 (2.1) は筆算をあらわしており、乗算の回数は  $ac, ad, bc, bd$  の4回である。一方、Karatsuba 法 (2.2) では加減算も合わせた演算の回数は増加しているが、乗算は  $ac, (a-b)(c-d), bd$  の3回に減少している。基底型の乗算の加減算に対する時間比を考えると、加減算の増加よりも乗算の回数の減少のほうが全体としての演算時間の減少に貢献している。この Karatsuba のアルゴリズムにおける乗算の回数は多倍長数を表す精度 (基底型の要素数)  $n$  に対して  $O(n^{\log_2 3}) \approx O(n^{1.585})$  となることが示され、計算量が  $O(n^2)$  である筆算に比べると、精度が大きいほど Karatsuba のアルゴリズムが優位になる。

他に代表的なものとしては、Fourier 変換を使うものがある。これは、筆算を合成積であると考えする方法である。大きい精度を要求される場合、例えば暗号理論に代表されるような場面では一般的に利用されており、専用ハードウェアによる実装例もある。しかし数千桁程度の演算では基底型の乗算以外の部分のオーバーヘッドが大きい。

exflib においては、目的とする桁数が100桁から数千桁程度と多倍長精度の研究としては比較的小さいことから、筆算および Karatsuba のアルゴリズムを実装している。乗算のルーチンが呼ばれると、要求された桁数に応じて適切なほうを選択して実行する。例えば AMD64 での筆算と Karatsuba 法の速度を比較すると図5のようになる。図中、横軸は計算精度を10進法で、縦軸は計算時間をマイクロ秒で表しており、実線が Karatsuba の計算時間を、破線が筆算のアルゴリズムを示している。10000桁程度までの範囲では筆算が高速であるため、こちらを利用する。一方、AMD64 とほぼ同じ命令セットを採用していてもマイクロアーキテクチャが異なる EM64T では、筆算と Karatsuba 法の速度は図6のようになり、6000桁程度までは筆算で、それを越えた場合には Karatsuba 法を利用するのがよいと考えられる。一般的なパーソナルコンピュータである IA-32 ではこの速度差の逆転が2000桁程度でおこることが図7よりわかる。

これらの図はプロセサのマイクロアーキテクチャに大きく依存するため [5]、最適な速度を実現するには、本来は計算環境毎に切り替え精度を決定するのが理想的である。配布している exflib では数台で速度比を検証した結果をもとに切り換える精度を埋め込んで固定している。

加減算と同様、基底型の乗算についても繰上がりが発生する。すなわち  $n$  ビットの符号なし整数の乗算は  $2n$  ビットになり、C 言語などでは表現可能な最大の整数同士の積を得ることはできず、積の上位  $n$  ビットの情報を得るには別途のアルゴリズムが必要である。しかし、プロセサの内部では積の上位を取得する特別な命令が定義されていることが多く、それを利用することにより  $2n$  ビットの積を得ることができる。

## exflib の除算、その他のアルゴリズムについて

除算には Ozawa のアルゴリズムを採用している [18, 24]。組み込み関数については引数を適切な範囲に還元した後、Taylor 展開を利用して函数値の計算をおこなっている。還元が続く Taylor 展開では引数が適切な範囲にあると仮定して固定小数点演算をおこなう。この固定小数点演算の部

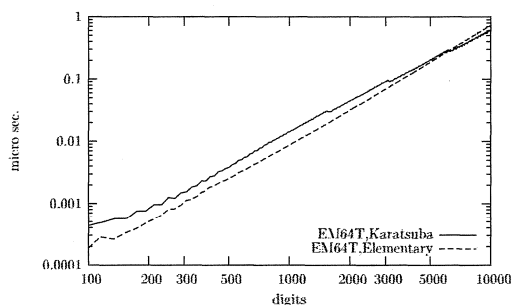
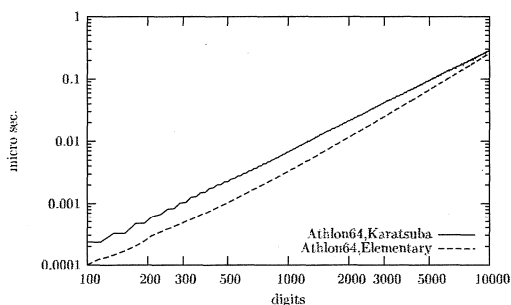


図 5: AMD64 での乗算の速度 (Athlon64 2.2GHz) 図 6: EM64T での乗算の速度 (Pentium4 3.6GHz)

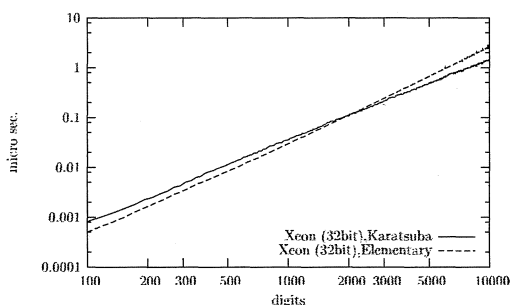


図 7: Xeon (32bit) での乗算の速度 (Xeon 3.2GHz)

分はアセンブリ言語で実装しており、それらを C 言語から呼び出す形で実装している。また展開に必要な項数の求め方は引数の範囲などによって異なるが、予め表で保存するか或は打ち切り誤差評価によって算出している。

## 丸め演算

演算の後には丸めをおこなう必要がある。exflib での丸めは、これらの演算を実現するサブルーチンの内部ではおこなわず、各サブルーチンは結果を 0 方向へ丸めた結果と丸めに必要な情報を返す。この演算サブルーチンに引き続き、丸め制御のサブルーチンと呼ぶことで適切に丸められた演算結果を得ることができる。exflib では四則に対する丸めは IEEE754 と同じく 4 つのモードを有しており、区間演算などへの利用も可能である。

## 引数渡し

演算を実現するサブルーチンの呼び出しにおいては、演算数、被演算数、演算結果を保持する変数、およびそれらの精度を引数として渡す。C 言語は値渡し (call by value) により引数をわたし、FOTRAN では参照渡し (call by reference) になる。exflib の四則演算ルーチンは、参照渡しのエントリと値渡しの両エントリを有しており、C 言語と FORTRAN のいずれからでも利用可能である。

## exflib の利用環境

現在までのところ、exflib は Alpha, IA-32, AMD64/EM64T のアーキテクチャを対象に実装している。アーキテクチャが限定であるのは、上述のように C 言語などでは利用できないプロセッサに固有の機能を多く利用しているためである。exflib では、アセンブリ言語で各演算サブルーチンを記述することでこれらの機能を利用している。また、C 言語などでは難しいソフトウェア・パイプラインニングやレイテンシの隠蔽などの技法を利用して高速化をはかっている [19]。入出力や初期化などは C 言語で実装している。

一般的な UNIX や PC 用の OS (Windows) などで利用可能である。現在までに UNIX では GCC, Windows では Visual C++, Borand C++, Cygwin, MinGW での動作を確認している。

アセンブリ言語で記述しているのは演算の部分のみである。このサブルーチンは上述のとおり C 言語と FORTRAN から利用可能である。入出力、メモリ管理やライブラリの内部で利用する数表、10 進法での入出力、初等関数は C 言語と C++ 言語で実装しており FORTRAN からは直接利用できない。

## C++ 言語用のインターフェース

exflib は C++ 言語からの利用を想定して実装しており、演算を実現する C 言語/FORTRAN 用のサブルーチンに加えて C++ 用のインターフェースを提供している。

C 言語において exflib ライブラリ中の演算サブルーチンを使用するためには、引数の型及び順序を考慮して適切な関数を呼び出す必要がある。具体例として多倍長数  $a, b, h$  に対して

$$S = \frac{(a+b)h}{2} \quad (2.3)$$

を計算する場合を考える。ユーザーは多倍長数型の一時変数 tmp1 および tmp2 を用意して

```
exflib_float_add(a, b, tmp1, precision);    // tmp1 = a + b
exflib_float_mul(tmp1, h, tmp2, precision); // tmp2 = tmp1 * h
exflib_float_idiv(tmp2, 2, S, precision);   // S = tmp2 / 2
```

と記述する必要がある。実際には丸め演算が必要であり、上記よりも複雑になるがここでは省略している。この例が意味するように、多倍長数を利用するプログラムの記述の際には、プログラマ自身が (2.3) の式を構文解析して 3 つの文を記述しなければならず、ユーザープログラムは非常に複雑なものとなり著しく可読性がおちる。さらに (2.3) の式を、多倍長数  $c$  に対して

$$S = \frac{(a+b)h}{c} \quad (2.4)$$

と変更する場合、プログラム中の最後の式を

```
exflib_float_idiv(tmp2, c, S);
```

と変更しただけでは、引数の型の不一致のためプログラムは正しく動作しない。つまりプログラムの変更に対してライブラリ単独のインターフェースは脆弱な環境であることがわかる。これを解決するためには、ポリモルフィズム (polymorphism) を利用すればよい。

ポリモルフィズムとは、複数のメソッド (method) に対して共通のインターフェースを提供することである。exflib では図 4 に示す多倍長数をクラス exfloat として定義し、そのクラスにおける演算子のオーバーロード (operator overloading) を利用してこれを実現している。これにより (2.3) の式の実現はプログラム中に

```
S = (a+b)*h/2;
```

とだけ記述すれば十分であり、(2.4) への変更に際しても

```
S = (a+b)*h/c;
```

によって、コンパイル時に型に応じて適切な関数が自動的に選択される。また、オーバーロードを利用しない場合には、各演算をおこなったあとに丸め演算を手動で呼び出す必要があるが、`exflib` ではオーバーロードにおいて、演算に引き続き丸めのためのサブルーチンの呼び出しもおこなっており、オーバーロードされた演算子を記述するだけでユーザが指定する丸め方向により適切な結果を得ることができる。

このようなポリモルフィズムによるインターフェースの利用により、オーバーヘッドは発生するものの、プログラムの負担の軽減、変更に対する強度の向上、型の不一致に対するセキュリティの向上、既存の C 言語プログラムの多倍長環境への移植性の向上などが期待される。

FOTRAN90 には、C++ 言語におけるクラス、オペレータオーバーロードに相当する機能がモジュールとして実装されている。今後 FORTRAN90 のモジュールを利用したインターフェースを実装することが課題のひとつである。

### 多倍長数の文字列による初期化

`exflib` では、演算子 `=` のオーバーロードを利用し、既存の整数型に加えて `NULL` 文字終端文字列 (C 言語の文字列) による多倍長数への代入を行うことが可能である。この文字列中には 10 進定数、 $\pi$  などの数学定数及びそれらの四則演算を記述することができる。具体的には `a`, `b` などを多倍長数型と宣言し

```
a = "0.1";  
b = 3;  
c = "2*#PI*3e2";  
d = "(-2.7 + 0.01828)/1.8";
```

と記述することができる。文字列中には円周率  $\pi$  と自然対数の底を含めることができる。演算子 `=` を実現するメソッドにより、要求した精度で多倍長数を代入することができる。文字列中に記述された式は、コンパイル時ではなく実行時に再帰下降解析 (recursive descent parsing) [1] を用いて構文解析を行っている。

### 出力書式の指定

`exflib` では、多倍長数の C++ 言語の出力ストリーム `ostream` への出力に際して、既存の型とのシームレスな環境を提供している。すなわち、書式指定には C++ 言語のマニピュレータ (manipulator) を用い、出力形式の指定には `iosetflags` を、また精度の指定には `setprecision` を利用する。これは倍精度方式と同じである。

### プログラム例

最後に、`exflib` を利用して指数関数  $e^x$  の Taylor 展開によって  $e^{\pi/2}$  を 10 進数で精度 100 桁で計算し、出力するためのプログラムの一例を示す。例に示すとおり、精度の指定は `exflib` で予約さ

れた PRECISION に代入することでおこなう。また、この例では示していないが、丸めモードの変更もクラス exfloat のメンバ変数でおこなう。

```
#define Precision 100
#include "exfloat.h"
using namespace std;

const int MAX_ITERATION = 100;

exfloat user_defined_exp(const exfloat &x)
{
    exfloat exp_x, term, old_value;

    exp_x = term = 1.0;
    old_value = 0.0;

    for(int n = 1; old_value != exp_x; n++)
    {
        if(n == MAX_ITERATION){
            cout << "Not converge." << endl;
            exit(1);
        }
        old_value = exp_x;
        term *= x/n;
        exp_x += term;
    }

    return exp_x;
}

int main()
{
    exfloat a = "#PI/2";

    cout << setprecision(100)
          << user_defined_exp(a)
          << endl;

    return 0;
}
```

#### 今後の課題

exflib は C++ 言語での利用を想定して設計と実装をつづけており、メモリ管理や入出力の部分は FORTRAN に充分に対応していない。これを FOTRAN に対応させ、同時に FORTRAN90 のモジュール機能を利用したインターフェースの提供も考えている。また、SPRAC V9, POWER5 などのアーキテクチャへの実装も進めている。

### 3 Chebyshev 多項式と函数近似

逆問題などの非適切問題の高精度な数値計算の実現には、多倍長数に代表される高精度な実数の近似とならび、微分方程式の高精度離散化も重要である。ここでは、数値流体力学の分野で利用されているスペクトル法について述べる。

スペクトル法では問題に応じて、Fourier 級数、Legendre 多項式、Chebyshev 多項式などにより試行函数を構成する。本稿では Chebyshev 多項式を利用する方法を述べる。

#### 記号の準備

$N$  を 3 以上の整数とする。  $0 \leq j \leq N$  に対して、

$$\bar{x}_j = \cos \frac{j\pi}{N}$$

を考える。スペクトル法ではこの  $\{\bar{x}_j\}$  を **Gauss-Lobatto (GL) 点** とよぶ。Chebyshev 点や Chebyshev-Gauss-Lobatto 点ともよばれる。

$$\bar{c}_j = \begin{cases} 2, & j = 0 \text{ または } N; \\ 1, & 1 \leq j \leq N-1 \end{cases}$$

とする。  $\delta_{ij}$  を Kronecker のデルタとする。区間  $[-1, 1]$  において  $k$  次の **Chebyshev 多項式**を

$$T_k(\cos \theta) = \cos k\theta, \quad k \geq 0$$

で定義する。  $T_0(x) = 1, T_1(x) = x$  であり、三角函数の加法定理より

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x) \quad (3.1)$$

である。区間  $[-1, 1]$  の両端において  $T_k(\pm 1) = (\pm 1)^k$ ,  $T'_k(\pm 1) = (\pm 1)^{k+1}k^2$  である。また

$$T_{k-1}(x) = \left\{ x + \frac{1}{k}(1-x^2)\frac{d}{dx} \right\} T_k(x) \quad (3.2)$$

が成立する。

Chebyshev 多項式は、  $L^2((-1, 1), dx/\sqrt{1-x^2})$  で直交する。Christoffel-Darboux の等式から、

$$\sum_{k=0}^{N-1} \frac{T_k(x)T_k(y)}{\bar{c}_k} = \frac{T_N(x)T_{N-1}(y) - T_{N-1}(x)T_N(y)}{x-y}, \quad x \neq y \quad (3.3)$$

であり、これから次が導かれる。

$$\sum_{k=0}^N \frac{T_k(\bar{x}_j)T_k(x)}{\bar{c}_k} = \frac{(1-x^2)T'_N(x)(-1)^{j+1}}{2N(x-\bar{x}_j)}, \quad 0 \leq j \leq N. \quad (3.4)$$

函数  $f(x)$  に対し、相異なる  $N+1$  点  $x_0, \dots, x_N$  上での  $f(x_0), \dots, f(x_N)$  の値が与えられているとする。このように、ある函数を代表させる点  $\{x_j\}$  を標本点 (sampling point) という。さて、  $f(x)$  を  $N$  次の多項式で近似することを考える。近似された多項式  $P_N f(x)$  は標本点上で  $f(x)$  の値に一致する、すなわち

$$P_N f(x_i) = f(x_i), \quad 0 \leq i \leq N$$

を満たすものとする.

$$Q_j(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{j-1})(x-x_{j+1})\cdots(x-x_N)}{(x_j-x_0)(x_j-x_1)\cdots(x_j-x_{j-1})(x_j-x_{j+1})\cdots(x_j-x_N)}$$

とする.  $Q_j(x)$  は  $N$  次の多項式で, 標本点  $\{x_i\}$  に対して

$$Q_j(x_i) = \delta_{ij}, \quad 0 \leq i, j \leq N$$

をみたす. したがって  $P_N f(x)$  は  $\{Q_j(x)\}$  をもちいて

$$P_N f(x) = \sum_{j=0}^N f(x_j) Q_j(x)$$

とすればよいことがわかる.  $\{Q_j(x)\}$  を標本点  $\{x_j\}$  に関する **Lagrange 補間多項式**,  $P_N f(x)$  を  $\{x_j\}$  に関する **Lagrange 補間の公式**という.

#### Gauss-Lobatto 点を標本点とする Lagrange 補間

区間  $[-1, 1]$  上の函数  $g_j(x)$  ( $0 \leq j \leq N$ ) を

$$g_j(x) = \frac{(1-x^2)T'_N(x)(-1)^{j+1}}{\bar{c}_j N^2(x-\bar{x}_j)}$$

と定義する. 式 (3.2) より

$$\begin{aligned} \frac{(1-\bar{x}_j^2)T'_N(\bar{x}_j)}{N} &= T_{N-1}(\bar{x}_j) - \bar{x}_j T_N(\bar{x}_j) \\ &= \cos \frac{j(N-1)\pi}{N} - \cos \frac{j\pi}{N} \cos \frac{jN\pi}{N} = 0 \end{aligned}$$

である.  $T'_N(x)$  は  $N-1$  次の多項式であり,  $j \neq 0, N$  のとき  $\bar{x}_j^2 \neq 1$  であるから,  $T'_N(x)$  は  $(x-\bar{x}_j)$  ( $1 \leq j \leq N-1$ ) を因子にもつ. 式 (3.1) から最高次の係数は  $2^{N-1}N$  であるから, 結局,

$$T'_N(x) = 2^{N-1}N(x-\bar{x}_1)(x-\bar{x}_2)\cdots(x-\bar{x}_{N-1})$$

となる.  $\bar{x}_0 = 1, \bar{x}_N = -1$  であることをあわせると,  $0 \leq j \leq N$  に対して  $g_j(x)$  は  $N$  次の多項式であることがわかる. また,

$$g_j(\bar{x}_i) = \delta_{ij}, \quad 0 \leq i, j \leq N \quad (3.5)$$

が示される. これは  $\{g_j(x)\}$  が GL 点に対する Lagrange 補間多項式であることを示しており, 次の定理が得られた.

**定理 1**  $u(x)$  の Gauss-Lobatto 点についての Lagrange 補間は

$$P_N u(x) = \sum_{j=0}^N u(\bar{x}_j) g_j(x) \quad (3.6)$$

である. すなわち  $P_N u$  は  $N$  次の多項式であり,  $P_N u(\bar{x}_i) = u(\bar{x}_i)$  が成立する.  $\square$

本稿では  $u_N(x) = P_N u(x), u_j = u(\bar{x}_j)$  と書くことがある.



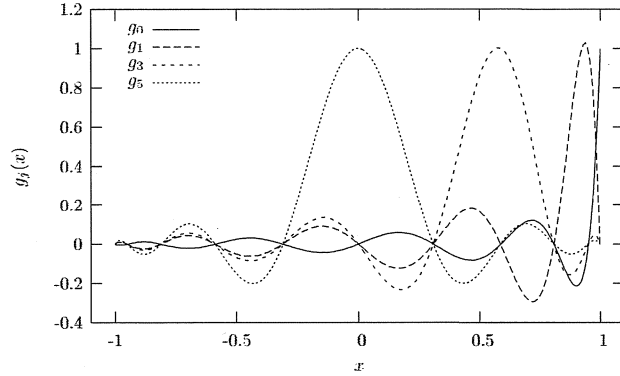


図 8:  $g_j(x)$  のグラフ ( $N = 10$  の場合)

### 形式的 Chebyshev 展開による表現

式 (3.4) より  $g_j(x)$  は

$$g_j(x) = \frac{2}{N\bar{c}_j} \sum_{k=0}^N \frac{1}{\bar{c}_k} T_k(\bar{x}_j) T_k(x)$$

となる。これをもちいると (3.6) は次のように書き直される。

$$\begin{aligned} P_N u(x) &= \sum_{j=0}^N u_j \frac{2}{N\bar{c}_j} \sum_{k=0}^N \frac{1}{\bar{c}_k} T_k(\bar{x}_j) T_k(x) \\ &= \sum_{k=0}^N \left\{ \frac{2}{N\bar{c}_k} \sum_{j=0}^N \frac{1}{\bar{c}_j} u_j T_k(\bar{x}_j) \right\} T_k(x). \end{aligned}$$

ここで  $0 \leq l \leq N$  に対して

$$\bar{a}_k = \frac{2}{N\bar{c}_k} \sum_{j=0}^N \frac{1}{\bar{c}_j} T_k(\bar{x}_j) u_j \quad (3.7)$$

とすると,  $\{u_j\}$  と  $\{\bar{a}_k\}$  は

$$\begin{pmatrix} T_0(\bar{x}_0) & T_1(\bar{x}_0) & \cdots & T_N(\bar{x}_0) \\ T_0(\bar{x}_1) & T_1(\bar{x}_1) & \cdots & \\ \vdots & & \ddots & \\ T_0(\bar{x}_N) & & & T_N(\bar{x}_N) \end{pmatrix} \begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \vdots \\ \bar{a}_N \end{pmatrix} = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_N \end{pmatrix} \quad (3.8)$$

を満たし, 行列  $(T_j(\bar{x}_i))$  の逆行列は (3.7) で与えられる。

この  $\{\bar{a}_k\}$  を利用すると, 函数  $u(x)$  の GL 点を標本点とする Lagrange 補間 (3.6) は, 次の形式的な Chebyshev 展開に書き換えられる。

$$P_N u(x) = \sum_{k=0}^N \bar{a}_k T_k(x). \quad (3.9)$$

逆に, (3.9) の右辺の形式的な Chebyshev 展開で与えられた函数は,  $\{g_j(x)\}$  をもちいて (3.6) の右辺の形に書き直すことができる。

#### 数値解析の視点から：補間の誤差評価

Lagrange 補間の射影作用素  $P_N$  の誤差評価が知られている。ここでは詳しく述べないが、 $u$  が解析的か、あるいは重みつき Sobolev 空間の意味で滑らかなとき、 $P_N u$  は、微分も込めて  $u$  の近似を与えることが知られている [3, 22].

#### 4 Chebyshev 多項式に基づくスペクトル法と微分方程式の数値計算

前節で示したとおり、Chebyshev 多項式による函数の形式的な展開は  $\{g_j(x)\}$  の線型結合に書き直される。微分方程式の離散化方程式は  $\{g_j(x)\}$  をもちいて簡潔に表現することができる。

##### Lagrange 補間の導函数

GL 点についての Lagrange 補間  $P_N u(x) = u_N(x)$  に対して、(3.6) より

$$\frac{du_N}{dx}(x) = \sum_{j=0}^N u_j g'_j(x)$$

である。GL 点  $\{\bar{x}_i\}$  においては

$$\frac{du_N}{dx}(\bar{x}_i) = \sum_{j=0}^N u_j g'_j(\bar{x}_i) \quad (4.1)$$

である。ここで、 $d_{ij} = g'_j(\bar{x}_i)$  とすると、 $d_{ij}$  の値は

$$\begin{aligned} d_{ij} &= \frac{\bar{c}_i}{\bar{c}_j} \frac{(-1)^{i+j}}{\bar{x}_i - \bar{x}_j}, & i \neq j, \\ d_{jj} &= -\frac{\bar{x}_j}{2(1 - \bar{x}_j^2)}, & j \neq 0, N, \\ d_{00} &= -d_{NN} = \frac{2N^2 + 1}{6} \end{aligned}$$

と求まり、(4.1) は次のように表される。

$$\frac{du_N}{dx}(\bar{x}_i) = \sum_{j=0}^N d_{ij} u_j. \quad (4.2)$$

微係数を  $u_j$  ( $0 \leq j \leq N$ ) の線型結合で表していることに注意する。さらに、

$$g'_j(x) = \sum_{l=0}^N g'_j(x_l) g_l(x) = P_N g'_j(x)$$

より  $g'_j(x)$  自身も Lagrange 補間になっているので、(4.2) より

$$g''_j(\bar{x}_i) = \frac{dg'_j}{dx}(\bar{x}_i) = \sum_{l=0}^N d_{il} (g'_j(\bar{x}_l)) = \sum_{l=0}^N d_{il} d_{lj} \quad (4.3)$$

となる。これをもちいると、 $u_N$  の 2 階導函数の値は次のようになる。

$$\frac{d^2 u_N}{dx^2}(\bar{x}_i) = \sum_{j=0}^N u_j g''_j(\bar{x}_i) = \sum_{j=0}^N \left( \sum_{l=0}^N d_{il} d_{lj} \right) u_j.$$

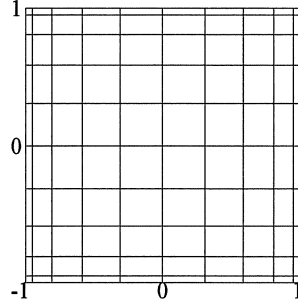


図 9:  $[-1, 1] \times [-1, 1]$  に配した GL 点 ( $N = 10$ )

高階微分も同様に表すことができる.

2変数の場合も同様である. 閉区間塊  $[-1, 1] \times [-1, 1]$  上の函数  $u(x, y)$  を考える.  $[-1, 1] \times [-1, 1]$  に GL 格子点  $\{(\bar{x}_m, \bar{y}_n)\}_{0 \leq m, n \leq N}$  を配し (図 9), この格子点上での函数  $u$  の値を  $u(\bar{x}_m, \bar{y}_n) = u_{mn}$  と書くことにする. 函数  $u$  の Lagrange 補間は次のようになる.

$$u_N(x, y) = P_N u(x, y) = \sum_{m, n=0}^N u_{mn} g_m(x) g_n(y).$$

この  $P_N u(x, y)$  の偏導函数は,

$$\begin{aligned} \frac{\partial u_N}{\partial x}(x, y) &= \sum_{m, n=0}^N u_{mn} g'_m(x) g_n(y), \\ \frac{\partial u_N}{\partial y}(x, y) &= \sum_{m, n=0}^N u_{mn} g_m(x) g'_n(y) \end{aligned}$$

であり, GL 点における値は (3.5) より

$$\begin{aligned} \frac{\partial u_N}{\partial x}(\bar{x}_i, \bar{y}_j) &= \sum_{m=0}^N u_{mj} g'_m(\bar{x}_i) = \sum_{m=0}^N d_{im} u_{mj}, \\ \frac{\partial u_N}{\partial y}(\bar{x}_i, \bar{y}_j) &= \sum_{n=0}^N u_{in} g'_n(\bar{y}_j) = \sum_{n=0}^N d_{jn} u_{in}. \end{aligned}$$

高階導函数の値も同様に, 次のとおり計算される.

$$\begin{aligned} \frac{\partial^2 u_N}{\partial x^2}(\bar{x}_i, \bar{y}_j) &= \sum_{m, n=0}^N u_{mn} g''_m(\bar{x}_i) g_n(\bar{y}_j) = \sum_{m=0}^N \left( \sum_{l=0}^N d_{il} d_{lm} \right) u_{mj}, \\ \frac{\partial^2 u_N}{\partial x \partial y}(\bar{x}_i, \bar{y}_j) &= \sum_{m, n=0}^N u_{mn} g'_m(\bar{x}_i) g'_n(\bar{y}_j) = \sum_{m, n=0}^N d_{im} d_{jn} u_{mn}, \\ \frac{\partial^2 u_N}{\partial y^2}(\bar{x}_i, \bar{y}_j) &= \sum_{m, n=0}^N u_{mn} g_m(\bar{x}_i) g''_n(\bar{y}_j) = \sum_{n=0}^N \left( \sum_{l=0}^N d_{jl} d_{ln} \right) u_{in}. \end{aligned}$$

## 微分方程式の離散化

スペクトル法による微分方程式の離散化には, pseudospectral 法, Galerkin 法, タウ法の 3 つがある. ここでは 2 点境界値問題

$$\frac{d^2 u}{dx^2}(x) = f(x), \quad -1 < x < 1, \quad (4.4)$$

$$u(1) = u(-1) = 0 \quad (4.5)$$

を例にして, Chebyshev 多項式で試行函数を表現する場合の離散化を説明する. 一般的な場合については [2, 9] に詳しい.

### pseudospectral method (スペクトル選点法)

未知函数  $u(x)$  を, Lagrange 補間 (3.6) にならって次で近似する.

$$u_N(x) = \sum_{j=0}^N \bar{u}_j g_j(x)$$

未知数は  $\{\bar{u}_j\}_{j=0}^N$  である. 区間  $(-1, 1)$  に, 相異なる  $N-1$  個の点  $\{x_i\}_{i=1}^{N-1}$  をとり,

$$\frac{d^2 u_N}{dx^2}(x_i) = \sum_{j=0}^N \bar{u}_j g_j''(x_i) = f(x_i), \quad 1 \leq i \leq N-1, \quad (4.6)$$

$$u_N(1) = \sum_{j=0}^N \bar{u}_j g_j(\bar{x}_0) = 0,$$

$$u_N(-1) = \sum_{j=0}^N \bar{u}_j g_j(\bar{x}_N) = 0$$

が成立するように  $\{\bar{u}_j\}$  を決定する数値解法を **Chebyshev 多項式による pseudospectral 法**, あるいは **Chebyshev 多項式によるスペクトル選点法 (spectral collocation method)** という. これは, 区間  $(-1, 1)$  で成立する微分方程式に対し, 有限個選んだ点で方程式の成立を要請する離散化手法である. 点  $x_i$  を選点 (collocation point) という.

選点  $\{x_i\}$  を GL 点からとる必要はない. 選点を GL 点  $\{\bar{x}_i\}_{i=1}^{N-1}$  とすると, (4.6) の  $g_j''(x_i)$  は (4.3) と書いて, 計算の手間を節約することができる.

### Galerkin 法

区間  $[-1, 1]$  上の  $N$  次多項式  $\mathbb{P}_N$  の中で, 境界条件 (4.5) をみたすものの集合を  $X_N$  とする.  $X_N$  の基底を一組とり,  $\{\phi_i\}_{i \in \Lambda}$  とする.  $\Lambda (= \Lambda_N)$  は添字の集合である. Lagrange 補間の方法で境界条件 (4.5) を満たす基底としては, 例えば

$$\phi_i = g_i, \quad i \in \Lambda = \{i; 1 \leq i \leq N-1\},$$

あるいは Chebyshev 多項式をもちいて

$$\phi_i = \begin{cases} T_i - T_0, & i : \text{even}; \\ T_i - T_1, & i : \text{odd}, \end{cases} \quad i \in \Lambda = \{i; 2 \leq i \leq N\}$$

とすればよい.

$$(u, v)_w = \int_{-1}^1 u(x) v(x) \frac{dx}{\sqrt{1-x^2}} \quad (4.7)$$

とする. 式 (4.4)–(4.5) に対して

$$\left( \frac{d^2 u_N}{dx^2}, v \right)_w = (f, v)_w, \quad v \in X_N$$

を満たす  $u_N \in X_N$  を求める解法を **Chebyshev 多項式による Galerkin 法** という. これは,  $X_N$  の基底に対して

$$\left( \frac{d^2 u_N}{dx^2}, \phi_i \right)_w = (f, \phi_i)_w, \quad i \in \Lambda \quad (4.8)$$

を満たす  $u_N \in X_N$  を求めることと同値である.

$$u_N(x) = \sum_{j \in \Lambda} \alpha_j \phi_j(x) \in X_N$$

とすると, Galerkin 法 (4.8) は, 連立一次方程式

$$\sum_{j \in \Lambda} \alpha_j \left( \int_{-1}^1 \phi_j''(x) \phi_i(x) \frac{dx}{\sqrt{1-x^2}} \right) = \int_{-1}^1 f(x) \phi_i(x) \frac{dx}{\sqrt{1-x^2}}, \quad i \in \Lambda \quad (4.9)$$

に帰着される.

#### 重み $1/\sqrt{1-x^2}$ を含む定積分の近似則

Galerkin 法に現れる (4.9) の定積分の計算には, 種々の方法がある. 例えば, 次の積分則がある.

$$\int_{-1}^1 p(x) \frac{dx}{\sqrt{1-x^2}} \approx \frac{\pi}{N} \sum_{l=0}^N \frac{p(\bar{x}_l)}{\bar{c}_l}. \quad (4.10)$$

被積分函数  $p(x)$  が十分に滑らかであれば, (4.10) の右辺は十分な精度の近似値を与える. 特に  $p(x)$  が  $2N-1$  次の多項式のとき, (4.10) の両辺は一致する (Lobatto-Markov の積分則 [21]). この積分公式 (4.10) の適用により, (4.9) の解  $\{\alpha_j\}$  は

$$\sum_{j \in \Lambda} \hat{\alpha}_j \left( \frac{\pi}{N} \sum_{l=0}^N \frac{1}{\bar{c}_l} \phi_j''(\bar{x}_l) \phi_i(\bar{x}_l) \right) = \frac{\pi}{N} \sum_{l=0}^N \frac{1}{\bar{c}_l} f(\bar{x}_l) \phi_i(\bar{x}_l), \quad i \in \Lambda \quad (4.11)$$

の解  $\{\hat{\alpha}_j\}$  によって近似される.

2 点, 注意を与える. 積分則 (4.10) を (4.9) に適用する際の離散化数を, 微分方程式の離散化数  $N$  に一致させる必要はない. また, 境界条件 (4.5) を満たす基底として  $\phi_i = g_i$  ( $1 \leq i \leq N-1$ ) を採ると, (3.5), (4.3) により (4.11) の係数の計算を節約することができる.

#### タウ法

Galerkin 法における基底  $\{\phi_i\}$  を境界条件をみたすように構成するのは, 境界条件が簡単な場合を除き, 容易ではない. このような場合には, タウ法がもちいられる.

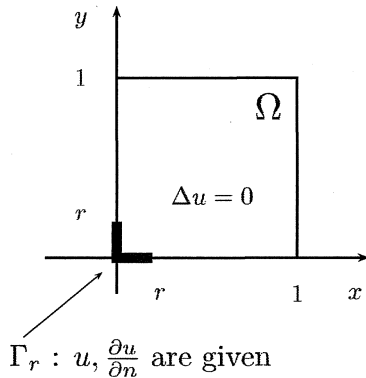


図 10: 正方形領域とデータ

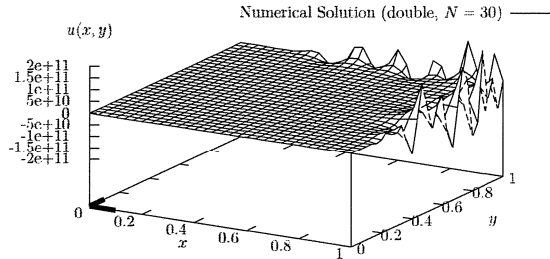


図 11: 正方形領域での数値解 (倍精度,  $N = 30$ )

### 数値解析の視点から：数値解の収束性

微分方程式のスペクトル解法の収束性について、数値解析の視点から述べておく。順問題に対しては安定性を利用した収束性の議論がある (例えば [11, 10])。しかし逆問題では安定性が成立することは期待されず、新たな数値解析理論が必要である。例えば差分法に対しては、早川 [12], 西田-磯 [23] らが、Kowalevskian を対象として、系の安定性を仮定せずに解析性をもちいた収束定理を与え、収束性と安定性が独立概念であることに改めて注意を与えた。安定性を仮定しない枠組でのスペクトル法の収束性は、著者の知る限り、未だ証明されていない。

## 5 数値計算例

多倍長計算と選点法により、逆問題と関連のある問題の数値計算例を示す。ここで紹介する例は、数学的には解析接続と関係のあるものである。

$\mathbb{R}^2$  の有界領域  $\Omega$  でラプラス方程式  $\Delta u = 0$  を満たし、 $\Omega$  の境界の一部  $\Gamma$  上で  $u(x, y) = f(x, y)$ ,  $\frac{\partial u}{\partial n}(x, y) = g(x, y)$  を満たす未知関数  $u(x, y)$  を求める問題を考える。ここで  $n$  は  $\Gamma$  における外向き単位法線とする。

まず  $\Omega$  を矩形領域  $\Omega = (0, 1) \times (0, 1)$  とし、 $\Gamma_r = \partial\Omega \cap \{x \leq 0.1, y \leq 0.1\}$  で  $f, g$  を与える場合を考える (図 10)。Chebyshev 多項式による pseudospectral 法をもちいて離散化することにより、問題の数値計算は連立一次方程式に帰着される。厳密解を  $u(x, y) = \frac{1}{1000\pi^2} \sin(3\pi x) \sinh(3\pi y)$  とし、スペクトル次数を  $N = 30$  として倍精度計算でおこなうと図 11 の数値解を得る。データを与える原点近傍と対角の頂点で数値解が振動しているのがわかる。一方、同一の設定での多倍長計算 (120 桁) では、図 12 の数値解を得る。この数値解の厳密解に対する最大誤差は  $2.0 \times 10^{-4}$  であり、必要な計算時間はパーソナルコンピュータ (Athlon64 3200+, Linux, gcc) で約 80 秒であった。さらにスペクトル次数  $N$  と数値誤差は図 5 のようになり、 $N$  に対して数値誤差は指数的に減衰していることがわかる。

次に、円領域  $\Omega = \{x^2 + y^2 < 1\}$  とその境界の一部  $\Gamma = \{(\cos \theta, \sin \theta); 0 \leq \theta \leq \pi/4\}$  の場合を考える (図 14)。厳密解  $u(x, y) = x^2 - y^2$  の設定に対し、120 桁、 $N = 30$  で得られる数値解を図 15 に示す。このとき、最大誤差は  $3 \times 10^{-12}$  未満であり、計算時間は約 80 秒であった。

最後に、円環領域  $\Omega = \{x^2 + y^2 < 1\} \setminus \{(x - 0.2)^2 + y^2 < 0.3^2\}$  と  $\Gamma = \{x^2 + y^2 = 1\}$  を考え (図 16)、厳密解を  $u(x, y) = \log \sqrt{(x - 0.15)^2 + (y - 0.1)^2}$  として数値計算をおこなう。この厳密

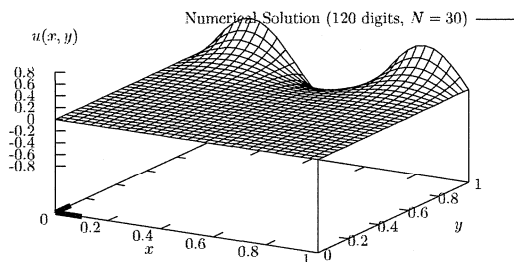


図 12: 正方領域での数値解 (120 桁,  $N = 30$ )

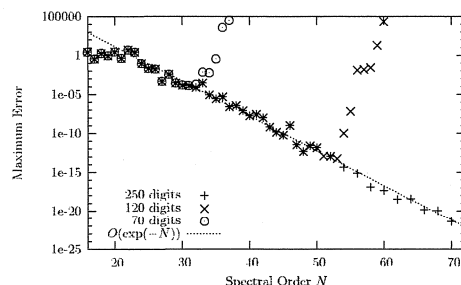


図 13: スペクトル選点法の数値解と厳密解との最大誤差

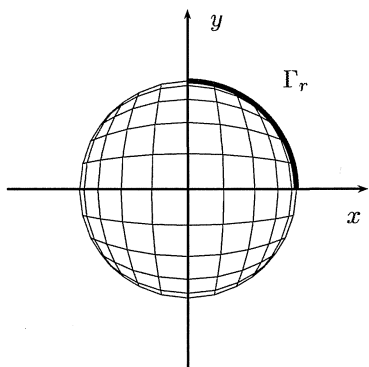


図 14: 円形領域

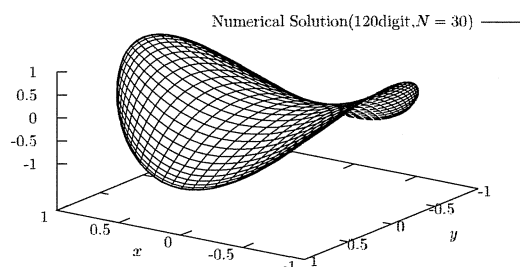


図 15: 単位円板での数値解 (120 桁,  $N = 30$ )

解  $u(x, y)$  は  $\Omega$  の外部に特異点をもち、この設定は心電図逆問題の設定と関連がある。この問題設定に対する数値解を図 17 に示す。数値解と厳密解との最大値誤差は、0.0015 未満であり、計算時間は約 73 秒であった。

いずれの例も数値的に不安定な問題設定であり、倍精度では実現し得なかった数値計算が多倍長計算により実現されている。このような高精度数値計算手法を利用することで数学解析理論を計算機上での展開が実現され、精密かつ信頼性の高い数値計算による逆問題へのアプローチが可能になるものと考えている。

## 参考文献

- [1] AHO, A. V., SETHI, R. and ULLMAN, J. D.: *Compilers, principles, techniques, and tools*. Addison-Wesley (1986).
- [2] CANUTO, C., HUSSAINI, M., QUARTERONI, A. and ZANG, T.: *Spectral Methods in Fluid Dynamics*. Springer-Verlag (1988).
- [3] CANUTO, C. and QUARTERONI, A.: Approximation results for orthogonal polynomials in Sobolev spaces. *Math. Comp.* **38** (1982), 67–86.

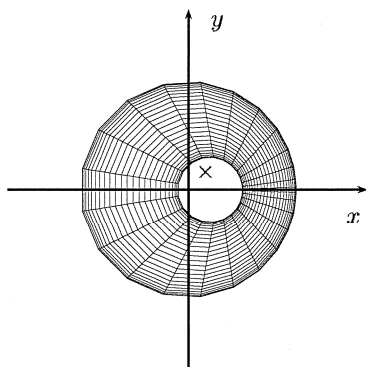


図 16: 円環領域,  $\times$  の位置  $(0.1, 0.15)$  が厳密解の特異点

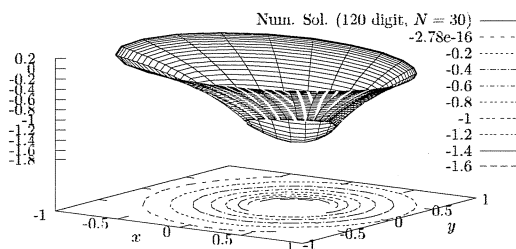


図 17: 円環領域での数値解と等高線 (120 桁,  $N = 30$ )

- [4] Compaq Computer Corporation: *Alpha Architecture Handbook* (1998). (Order Number: EC-QD2KC-TE).
- [5] Exflib benchmark: <http://www-an.acs.i.kyoto-u.ac.jp/~fujiwara/exflib/exflib-index-benchmark.html>.
- [6] FUJIWARA, H.: Numerical analysis of ill-posed problems based on multiple-precision arithmetic (2003). (博士学位論文).
- [7] GOLDBERG, D.: What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys* **23** (1991), 5–48.
- [8] GOLDBERG, D.: *Computer Arithmetic* (1996). Appendix in [20].
- [9] GOTTLIEB, D., HUSSAINI, M. Y. and ORSZAG, S. A.: Theory and applications of spectral methods. *Spectral Methods for Partial Differential Equations*, SIAM (1984), 1–54.
- [10] GOTTLIEB, D., LUSTMAN, L. and TADMOR, E.: Convergence of spectral methods for hyperbolic initial-boundary value systems. *SIAM J. Numer. Anal.* **24**(3) (1987), 532–537.
- [11] GOTTLIEB, D., LUSTMAN, L. and TADMOR, E.: Stability analysis of spectral methods for hyperbolic initial-boundary value systems. *SIAM J. Numer. Anal.* **24**(2) (1987), 241–256.
- [12] HAYAKAWA, K.: Convergence of finite difference scheme and analytic data. *Publ. Res. Inst. Math. Sci.* **24** (1988), 759–764.
- [13] HIGHAM, N. J.: *Accuracy and Stability of Numerical Algorithms*. SIAM (2002).
- [14] Intel Corporation: *IA-32 Intel® Architecture Software Developer's Manual* (2003).
- [15] ISO/IEC 9899 : 1999 - Programming Language C.
- [16] KNUTH, D. E.: *The Art of Computer Programming Volume 2 : Semi Numerical Algorithms*, 3rd ed. Addison-Wesley (1998).



- [17] MULLER, J.-M.: *Elementary Functions – Algorithms and Implementation*. Birkhäuser (1997).
- [18] OZAWA, K.: A fast  $O(n^2)$  division algorithm for multiple-precision floating-point numbers. *J. of Information Processing* **14** (1991), 354–356.
- [19] PATTERSON, D. A. and HENNESSY, J. L.: *Computer Organization and Design*. Morgan Kaufmann (1994).
- [20] PATTERSON, D. A. and HENNESSY, J. L.: *Computer Architecture: A Quantitative Approach*, 2nd ed. Morgan Kaufmann (1996).
- [21] RIVLIN, T. J.: *The Chebyshev Polynomials*. Wiley (1974).
- [22] TADMOR, E.: The exponential accuracy of Fourier and Chebyshev differencing methods. *SIAM J. Numer. Anal.* **23** (1982), 1–10.
- [23] 西田詩, 磯祐介: Kowalewsky の定理と差分法. 日本数学会 2005 年度年会 応用数学分科会講演アブストラクト (2005), 174–175.
- [24] 藤原宏志, 磯祐介: 64bit 計算環境に適した多倍長数値計算環境の構築と非適切問題の数値計算. 情報処理学会論文誌 **44** (2003), 925–931.